

**HOCHSCHULE
MITTWEIDA**
UNIVERSITY OF
APPLIED SCIENCES



Fachbereich Medien

Nebel, Steve

Erstellung eines Simulationstools für
Kommunikationsprozesse im Internet

Building a Simulation Tool for Communication
Processes on the Internet

- Bachelorarbeit -

Hochschule Mittweida – University of Applied Science (FH)

Mittweida - 2010

**HOCHSCHULE
MITTWEIDA**
UNIVERSITY OF
APPLIED SCIENCES



Fachbereich Medien

Nebel, Steve

Erstellung eines Simulationstools für
Kommunikationsprozesse im Internet

Building a Simulation Tool for Communication
Processes on the Internet

- eingereicht als Bachelorarbeit -
Hochschule Mittweida – University of Applied Science (FH)

Erstprüfer: Prof.Dr.-Ing.Robert J. Wierzbicki
Zweitprüfer: Dr. Lothar Krempel

Mittweida – 2010

Bibliographische Beschreibung

Nebel, Steve

Erstellung eines Simulationstools für
Kommunikationsprozesse im Internet - Building a
Simulation Tool for Communication Processes on
the Internet - 2010 – 71 Seiten, Mittweida,
Hochschule Mittweida(FH), Fachbereich Medien,
Bachelorarbeit

Referat

In dieser Bachelorarbeit wird die vom Autor der Arbeit durchgeführte Entwicklung eines Simulationstools für die Informationsausbreitung im Internet beschrieben. Dazu wird sich mit Theorien der Informationsausbreitung, der Netzwerkstruktur und der Simulationsmöglichkeiten befasst. Mit Hilfe dieser theoretisch-wissenschaftlichen Grundlagen wird der Entwicklungsprozess begründet und mit einer Beispielsimulation abgeschlossen.

Inhaltsverzeichnis

Abbildungsverzeichnis.....	IV
Tabellenverzeichnis.....	VI
1 Einleitung	1
1.1 Problemstellung	1
1.2 Zielsetzung	2
1.3 Vorgehensweise	3
2 Begriffsabgrenzung.....	5
2.1 Begriffsvielfalt	5
2.2 Zwei Klassen der Opinion Leader	6
2.3 Social Hubs.....	6
2.4 Knowledge Hubs.....	7
2.5 Innovators und Early Adopters	7
3 Netzwerkerstellung.....	9
3.1 Netlogo.....	9
3.2 Knotengradverteilung	9
3.2.1 Zufallsgraphen	9
3.2.2 Generalisierte Zufallsgraphen	10
3.2.3 Power Law.....	11
3.2.4 Preferential Attachment	12
3.2.5 Umsetzung des Netzwerkes in Netlogo.....	13
3.3 Ein Netzwerk voller Dreiecke	14
3.4 Mega Hubs.....	17
3.5 Dimensionierung des Knotengrades	19
3.5.1 Soziale Kontakte	19
3.5.2 Weitergabennetzwerk	20
3.5.3 Die Erstellung von Social Hubs	21
3.6 Obergrenze der Social Hubs.....	24
4 Simulationenmethoden.....	26

4.1	Verwendung des Netzwerks	26
4.2	Binäre Entscheidungen	26
4.3	Threshold Modell	27
4.3.1	Grundlagen des Threshold	27
4.3.2	Umsetzung des Threshold Modells in Netlogo	29
4.3.3	Wahl des Thresholds	30
4.3.4	Simulationsentscheidende Thresholdwerte	31
4.4	Die SI-, SIS- und SIR-Modelle	34
4.4.1	Auswahl des Modells für die Simulation	35
4.4.2	Umsetzung des SIR-Modells in Netlogo	36
4.5	Kombination der Modelle	37
5	Knotenbeschriftung	39
5.1	Zwei Knotenklassen	39
5.2	Seeding	40
5.3	Weitergabe der Innovators	41
6	Beispielsimulation	43
6.1	Ausgangslage	43
6.1.1	Netzwerkerstellung	43
6.1.2	Weitergabeparameter	44
6.2	Beispielseeding	45
6.3	Weitergabewahrscheinlichkeit	48
6.4	Annahmewahrscheinlichkeit	51
6.4.1	Knowledge Hubs zu Imitators	52
6.4.2	Knowledge Hubs zu Knowledge Hubs	53
6.4.3	Imitators zu Imitators	54
6.4.4	Imitators zu Knowledge Hubs	55
6.5	Netzwerkstruktur	55
6.6	Fazit der Beispielsimulation	58
7	Fazit	60
	Literaturverzeichnis	62
	Fachbücher	62
	Hochschulschriften	62

Sonstige Schriften	63
Zeitschriften	64
Anhang.....	65
Vorwort zu den Kurzanleitungen	65
Kurzanleitung für das Tool zur Netzwerkerstellung	66
Kurzanleitung für das Tool zur Netzwerkbeschriftung	67
Kurzanleitung für das Tool zur Simulationsdurchführung	69
Selbständigkeitserklärung	71

Abbildungsverzeichnis

Abbildung 1: Simulation der Knotengradverteilung eines Erdős-Rényi-Zufallsgraphen	10
Abbildung 2: Zufallsgraphen mit festgelegter Knotengradverteilung	11
Abbildung 3: Wachstum einer Kochkurve	12
Abbildung 4: Ausgangslage der Netzwerkerstellung	16
Abbildung 5: Zweiter Schritt der Netzwerkerstellung	16
Abbildung 6: Dritter Schritt der Netzwerkerstellung	16
Abbildung 7: Knotengradverteilung der aktuellen Netzwerkerstellung bei 3000 Knoten	17
Abbildung 8: Knotengradverteilung der aktuellen Netzwerkerstellung bei 3000 Knoten und einem Abweichwert von -0,05	18
Abbildung 9: Knotengradverteilung der aktuellen Netzwerkerstellung bei 3000 Knoten und einem Abweichwert von 0,05	18
Abbildung 10: Anzahl der Leute mit denen Internet Nutzer Inhalte teilen	21
Abbildung 11: Vergleich der Knotengrade des Simulationsnetzwerkes mit den Daten aus Abbildung 10	22
Abbildung 12: Erneuter Vergleich mit den Daten aus Abbildung 10	23
Abbildung 13: Kontengradverteilung vor CSH	23
Abbildung 14: Kontengradverteilung nach CSH	24
Abbildung 15: Versuchsbild des Asch-Experiments	28
Abbildung 16: Eine mögliche Thresholdverteilung	31
Abbildung 17: Informationsausbreitung mit einem Threshold von 0,33	32
Abbildung 18: Informationsausbreitung mit einem Threshold von 0,34	33
Abbildung 19: Informationsausbreitung mit einem Threshold von 0,66	33
Abbildung 20: Informationsausbreitung mit einem Threshold von 0,67	34
Abbildung 21: Beispiel einer SIR-Simulation	36
Abbildung 22: Zwei Simulationsdurchläufe mit Thresholdabschaltung (1)	38
Abbildung 23: Zwei Simulationsdurchläufe mit Thresholdabschaltung (2)	38
Abbildung 24: Anzahl der Early Adopter nach Seedingmethode	41
Abbildung 25: Knotengradverteilung des Beispielnetzwerkes	44
Abbildung 26: Informationsausbreitungen	45
Abbildung 27: Typischer Ausbreitungsverlauf der Beispielsimulation	46
Abbildung 28: Auswirkung der Weitergabewahrscheinlichkeit der Innovators	47

Abbildung 29: Auswirkungen der Weitergabewahrscheinlichkeit auf die Informationsausbreitung	49
Abbildung 30: Auswirkungen der Weitergabewahrscheinlichkeit auf die Ausbreitungsgröße	50
Abbildung 31: Auswirkungen der Weitergabewahrscheinlichkeit auf die Ausbreitungsdauer	50
Abbildung 32: Auswirkungen der Annahmewahrscheinlichkeit $KH \rightarrow IM$	52
Abbildung 33: Auswirkungen der Annahmewahrscheinlichkeit $KH \rightarrow KH$	53
Abbildung 34: Auswirkungen der Annahmewahrscheinlichkeit $IM \rightarrow IM$	54
Abbildung 35: Auswirkungen der Annahmewahrscheinlichkeit $IM \rightarrow KH$	55
Abbildung 36: Knotengradverteilung des Versuchsnetzwerkes 2	56
Abbildung 37: Vergleich mit Versuchsnetzwerk 2	56
Abbildung 38: Knotengradverteilung des Versuchsnetzwerkes 3	57
Abbildung 39: Vergleich mit Versuchsnetzwerk 3	58
Abbildung 40: Das Tool zur Netzwerkerstellung	66
Abbildung 41: Das Tool zur Netzwerkbeschriftung	67
Abbildung 42: Das Tool zur Simulationsdurchführung	69

Tabellenverzeichnis

Tabelle 1: Ausgangswerte der Beispielsimulation	45
Tabelle 2: Beispielseeding mit Einzelknoten	46
Tabelle 3: Weitergaben und Ausbreitungswahrscheinlichkeit im Vergleich ...	48

1 Einleitung

1.1 Problemstellung

In der Welt des Marketings haben Methoden wie virales Marketing, Empfehlungsmarketing oder auch Buzz Marketing in den letzten Jahren eine gewaltige Aufmerksamkeit erfahren. Dem Reiz mit geringem Aufwand eine exponentielle Verbreitung zu erreichen und zusätzlich, dank der zwischenmenschlichen Werbung, noch effektiver zu sein als mit klassischen Werbeformen, konnten sich wenige entziehen. Doch der tatsächliche Erfolg der Werbeansätze variiert stark. Hatten einige Kampagnen den beschriebenen sensationellen Erfolg, scheiterten andere völlig.

Dem folgt auch die wissenschaftliche Arbeit im Gebiet der Mundpropaganda. Aus der anfänglich einfachen Struktur und Methodik entstanden unterschiedliche Simulationsansätze und Kommunikationstypen. Es wird zunehmend ersichtlich, dass keine Zielgruppe mit der anderen übereinstimmt und einfache Verbreitungsansätze nicht universell anwendbar sind. Zusätzlich erschwert die stetige Evolution der Kommunikation im Internet die Bestimmung einheitlicher Grundsätze.

Um dennoch Aussagen über die Verbreitung und Wirksamkeit von Mundpropaganda treffen zu können, empfiehlt es sich eine flexible und anpassungsfähige Simulationslösung zu entwickeln. Mit dieser könnten neben allgemeinen Aussagen, auch genauere zielgruppenspezifische Werte, wie die Struktur des Netzwerkes oder Weitergabeparameter, Berücksichtigung finden. So kann überprüft werden, ob die gleichen Informationen sich in unterschiedlichen Netz- und Kommunikationsstrukturen verbreiten. Auch der umgekehrte Fall ist denkbar, wie sich unterschiedliche Informationen bei den gleichen Gegebenheiten ausbreiten.

Nur so erscheint es möglich die Simulationen auf ein Netzwerk genau anzupassen und somit relevante Ergebnisse zu erzielen.

Ergebnisse, mit denen sich die Verbreitung und Wirksamkeit der Informationsausbreitung untersuchen und optimieren lässt.

1.2 Zielsetzung

Es wird ein Netzwerk konstruiert mit dessen Hilfe sich das Verhalten sozialer Netzwerke simulieren lässt. Anschließend wird dieses zur Simulation von Informationsausbreitung genutzt.¹ Dabei findet die Umgebung Netlogo Verwendung. Eine Programmierumgebung, die am "Northwestern University's Center for Connected Learning" entwickelt wird² und Agent Based Models verwendet.³ (ABM)

Die Arbeit mit ABM ermöglicht neue Einsichten in den Ablauf komplexer und dynamischer Prozesse. Solche Prozesse sind nur schwer zu erfassen und exakt zu beschreiben. Dies gilt auch für die Informationsausbreitung. Fragen, wie beispielsweise wann treten überdurchschnittlich große Ausbreitungen auf oder welche Kommunikatoren sind besonders wichtig, lassen sich nur schwer klären.

Mit Hilfe der ABM kann ein System von Agenten erstellt werden, welches klaren aber flexibel einstellbaren Regeln folgt. Aus dem Verhalten der einzelnen Agenten in der Mikro Ebene⁴ entwickelt sich ein Ablauf der Informationsverbreitung auf der Makro Ebene⁵, der erfasst und ausgewertet werden kann. Diese Methode gewährt Einblicke in das komplexe Zusammenspiel der Kommunikatoren.

Ein weiterer Nutzen, der aus der Modellierung mit Hilfe der ABM gezogen werden kann, ist das Erkennen von emergenten Effekten⁶ in der Informationsausbreitung, beispielsweise von Schwellenwerten im Ausbreitungsverlauf ab denen sich die

¹ Die Simulation von Informationsausbreitung kann analog zur Simulation von "Word of Mouth" betrachtet werden, denn "Word of Mouth" ist nur ein anderer Begriff für die interpersonelle Informationsweitergabe.

² vgl. Wilensky 1999

³ Deutsch: auf Agenten basierende Modelle

⁴ Diese Ebene bezeichnet einzelne Agenten

⁵ Diese Ebene bezeichnet das gesamte Netzwerk

⁶ Emergente Effekte sind Effekte, die auf der Makro Ebene durch das Zusammenspiel der einzelnen Agenten in der Mikro Ebene entstehen.

Informationen ausbreiten, ohne dass diese in den Regeln der Agenten vordefiniert wurden.

Neben den Vorteilen, die sich bei der Modellierung mit ABM ergeben, werden auch deren Einschränkungen bei der Zielsetzung bedacht. ABM sind nicht so konstant wie rein mathematische Herangehensweisen, d.h. sie folgen keinen fest definierten Funktionen und Abläufen, sondern entfalten sich dynamisch, im schlimmsten Fall sogar chaotisch.⁷ Diese Unbestimmbarkeit hat den Nachteil, dass sich kaum perfekte Definitionen formulieren lassen, aber auch den Vorteil, dass bedeutend mehr Möglichkeiten und Abläufe untersucht werden können.

So kann das Ziel des erstellten Modells nicht sein, eine Informationsverbreitung im Internet exakt vorherzusagen, wohl aber aus den Eigenschaften Einsicht in die Funktionsweise zu gewinnen und Rückschlüsse auf Verbesserungsmöglichkeiten zu ziehen.

Somit eignet sie die vorliegende Arbeit für Hochschulen oder Studenten, die mittels des Tools die Funktionsweise und Eigenschaften der Informationsausbreitung verdeutlichen wollen. Genauso geeignet ist das Tool für die wissenschaftliche Simulation, sowie für eine Betrachtung eines konkreten Sachverhaltes aus der wirtschaftlichen Realität.

Die Arbeit eignet sich auch für jene, die an einem Überblick über den aktuellen Informationsstand zu den Mitgliedern und der Struktur von Kommunikationsnetzwerken interessiert sind. Sie erlaubt zusätzlich einen Einblick in die Simulation der Ausbreitung von Informationen in sozialen Netzen.

1.3 Vorgehensweise

In der vorliegenden Arbeit wird zunächst die Umsetzung des Kommunikationsmodells Schritt für Schritt erläutert und begründet. Dafür werden im Kapitel 2 zunächst die benötigten Begrifflichkeiten geklärt. Diese finden in der Literatur teilweise synonyme oder gegensätzliche Verwendung, was eine einheitliche Definition erschwert.

⁷ vgl. Page 2005

Danach wird sich im Kapitel 3 mit der Erstellung des Netzwerkes beschäftigt, mit dem Ziel eine Lösung zu finden, die den modernen Bedingungen menschlicher Kommunikation im Internet entsprechen kann. Gleichzeitig soll sie modifizierbar genug gestaltet sein, um sie an abgewandelte Formen anpassen zu können.

Im Kapitel 4 werden unterschiedliche Simulationsmethoden erläutert und miteinander verglichen, wobei der Fokus auf der Betrachtung der Grundlagen und der Umsetzung in Netlogo liegt.

Anschließend erfolgt im Kapitel 5 die Festlegung der einzelnen Bestandteile des Netzwerkes entsprechend ihrer Rolle bei der Verteilung. Dadurch wird das Netzwerk für die Simulation vorbereitet und bestimmt.

Im Kapitel 6 werden die Ergebnisse einer Beispielsimulation beschrieben. Dabei wird der Einfluss diverser Parameter auf die Simulationsergebnisse diskutiert.

Eine abschließende Bewertung des erstellten Tools und dessen Möglichkeiten ist Gegenstand des abschließenden Kapitel 7.

Um das komplette Vorgehen nachvollziehen zu können und eventuell weiter mit dem erstellten Tool zu arbeiten, sind geringe Vorkenntnisse im Bereich der Netzwerktheorie und Kommunikationsforschung sinnvoll, aber nicht unbedingt notwendig.

2 Begriffsabgrenzung

2.1 Begriffsvielfalt

Die Liste mit Begriffen, die Kommunikationsstruktur und Typen bei der Verbreitung von Informationen beschreiben sollen, ist umfangreich. Es existieren viele, in der Kommunikationsforschung häufig verwendete Begriffe wie, „Innovator“, „Vermittler“, „Verkäufer“, „Kenner“, "Opinion Leader", "Early Adopter", und "Imitators". Bei deren Verwendung kommt es in der Literatur zu Überschneidungen und Widersprüchen. Für eine exakte Nachvollziehbarkeit wird sich daher vor der Entwicklung einer Simulationslösung mit den benötigten Begriffen auseinander gesetzt.

Ein Beispiel für diese Überschneidungen ist der Begriff der "Opinion Leader". Diese, auch "Influentials" genannten Personen, verfügen über ein großes Produkt- und Informationswissen, haben eine größere Anzahl sozialer Kontakte als der Durchschnitt und sind innovativer und wagemutiger als andere.⁸ Malcolm Gladwell hingegen definiert hierzu die Kommunikationstypen des Vermittlers, der eine große Anzahl an sozialen Kontakten besitzt, des Kenners, der über ein besonders großes Produktwissen verfügt und des Verkäufers, der einen besonderen Einfluss auf seine Mitmenschen hat.⁹ Diese Überschneidung der Begriffe allein ist jedoch nicht problematisch. Bei Gladwell müssen die drei definierten Typen allerdings nicht in einer Person vereint sein. Dies widerspricht dem klassischen Begriff des Opinion Leaders, sodass er nicht synonym mit den Einzelbegriffen verwendet werden kann. Zusätzlich dazu verwendet Matthias Rürup den Begriff Opinion Leader synonym mit einer Schicht, die bei der Verbreitung von Informationen nur eine Vorreiterrolle einnehmen, die demzufolge nicht in der späteren Verbreitung vorkommen.¹⁰ Dies ist aber keine Bedingung aller vier bis jetzt genannter Definitionen.

An diesem Beispiel ist abzusehen, warum für eine Simulation die Begriffe genau festgelegt werden müssen. Denn die

⁸ vgl. Chan; Misra 1990, 55

⁹ vgl. Gladwell 2002, 41 - 105

¹⁰ vgl. Rürup 2007, 66

Ergebnisse der Simulation und die Modellierung müssen eindeutig und nachvollziehbar sein.

2.2 Zwei Klassen der Opinion Leader

Bereits Michaela Klinger ordnete die Begriffe, die sich mit den Eigenschaften der Opinion Leader beschäftigen in mehrere Gruppen ein.¹¹ Sie unterteilte die Meinungsführer in zwei Klassen. Zum einen die Experten-Meinungsführer, zu denen sich die Kenner von Gladwell zählen lassen. Zum anderen die gesellschaftlichen Meinungsführer, zu denen die Vermittler von Gladwell gehören. Bei dieser Einteilung taucht der Begriff des Verkäufers von Gladwell nicht mehr auf. Dies deckt sich mit den meisten Arbeiten zum Word of Mouth, bei denen nur zwischen zwei und nicht drei Kommunikationstypen unterschieden wird. Des Weiteren wird zwischen der Einbindung des Kommunikators in das Netzwerk und dessen Verhalten im Umgang mit Informationen unterschieden. Eine Unterteilung, die auch in weiteren Arbeiten Verwendung findet. Beispielsweise bei Emanuel Rosen, der die zwei Typen Social- und Expert Hubs verwendet.¹² Eine solche Unterteilung scheint auch für das Erstellen der Simulation sinnvoll, da die strukturellen Eigenschaften bei der Erstellung des Netzwerkes beachtet werden müssen und unabhängig davon das Verhalten bei der Durchführung der Simulation Beachtung findet.

2.3 Social Hubs

Der erste Begriff der für die Erstellung des Netzwerkes wichtig ist, ist der des Social Hubs. Er bezeichnet Kommunikatoren mit deutlich mehr Verbindungen und sozialen Kontakten als der Durchschnitt. Gladwells Begriff der Vermittler lässt sich hierzu synonym verwenden. Das besondere an Social Hubs ist, dass sich relativ sichere Werte zu ihrer Anzahl und zu ihren sozialen Verbindungen

¹¹ vgl. Klinger 2006, 86

¹² vgl. Rosen 2009, 83 - 107

Ein anderer Begriff für Experten Hubs ist Knowledge Hubs

bestimmen lassen. Durch Analysen sozialer Netzwerke und statistischer Erhebungen wurden bereits verschiedenste Werte ermittelt.¹³ Diese stimmen nicht immer überein, lassen aber Rückschlüsse auf die ungefähren Dimensionen zu. Dadurch können Social Hubs eine zentrale Rolle bei der Entwicklung des Simulationsnetzwerkes einnehmen.

2.4 Knowledge Hubs

Im Gegensatz zum strukturellen Begriff der Social Hubs, ist die Einbettung in das Netzwerk bei Knowledge Hubs nicht der ausschlaggebende Faktor. Knowledge Hubs verfügen über ein großes Informations- und Fachwissen. Sie unterscheiden sich in Weitergabe- und Annahmeverhalten von normalen Kommunikationsteilnehmern. Dabei ist zu beachten, dass Knowledge Hubs und Social Hubs durchaus in einer Person vereint sein können, es aber nicht zwangsläufig sein müssen. Daher wird in der restlichen Arbeit nicht mit dem klassischen Opinion Leader Begriff gearbeitet, dem man meistens die Eigenschaften eines Social und Knowledge Hubs unterstellt.¹⁴ Durch diese Trennung kann die Simulation auch deutlich flexibler erfolgen, da Fälle untersucht werden können in denen sich unter den Social Hubs viele Experten befinden, sowie Fälle in denen dies nicht der Fall ist.

2.5 Innovators und Early Adopters

Weitere Kommunikationstypen, die in der Literatur Verwendung finden, sind die Innovators und die Early Adopter. Diese basieren auf der Annahme von Everett Rogers, dass Menschen unterschiedlich auf Innovationen und Neuerungen reagieren.¹⁵ Innovators bezeichnen dabei die ersten 2,5% der Individuen, die eine Neuerung aufnehmen. Sie sind besonders innovations- und risikofreudig. Early Adopter bezeichnen die nächsten 13,5%, die eine Innovation

¹³ vgl. Viral Manager - How Users Share Viral Content o.J., 2

¹⁴ vgl. Hermann 2004, 30f

¹⁵ vgl. Rogers 1995

Begriffsabgrenzung

aufnehmen. Sie zeichnen sich durch ihre besondere Position und Vorbildfunktion im sozialen System aus.

Obwohl diese Begriffe nicht aus der Netzwerkforschung stammen, bieten sie sich aufgrund ihrer Eigenschaften für die Bezeichnung von bestimmten Informationsträgern in der Simulationslösung an. Dazu müssen sie leicht abgewandelt werden. Als Innovators werden im Folgenden die ersten manuell ausgewählten Informationsträger bezeichnet. Als Early Adopter wird die Anzahl der möglichen Informationsträger im ersten Schritt der Ausbreitung bezeichnet. Somit können die Vorreiter der Informationsausbreitung sinnvoll identifiziert werden. Diese Begriffsdefinition hat sich bereits bei anderen Simulationslösungen bewährt.¹⁶

¹⁶ vgl. Watts 2002, 2

3 Netzwerkerstellung

3.1 Netlogo

Zur Erstellung und Simulation des Netzwerkes wird die Programmierungsumgebung Netlogo verwendet. Damit ist es relativ einfach ein Netzwerk aus Knoten (Nodes) und Verbindungen (Links) zu erstellen. Für die Knoten können nahezu beliebig viele Eigenschaften definiert werden, und sie können frei untereinander verknüpft werden. Die einzelnen Knoten des Systems werden unabhängig behandelt, was sich im Fall einer Netzwerksimulation als besonders vorteilhaft erweist. So lässt sich relativ unkompliziert die Wechselwirkung zwischen den Knoten simulieren und auswerten.

3.2 Knotengradverteilung

Als erster Schritt zur Erstellung eines Netzwerkes wird sich mit der Knotengradverteilung beschäftigt. Als Knotengrad wird die Anzahl der Verbindungen (Kontakte), die ein Knoten (Kommunikationsteilnehmer) besitzt, bezeichnet. Hierfür gibt es in der Kommunikationsforschung zahlreiche Hinweise, die als Richtlinie dienen können.

3.2.1 Zufallsgraphen

Die einfachste Methode ein Netzwerk zu erstellen ist es, es rein zufällig zu entwickeln. Bei dieser Methode entstehen Netzwerke in denen die Knoten einen vorgegebenen durchschnittlichen Knotengrad besitzen. So lässt sich eine relativ gleichmäßige Knotengradverteilung erzielen. (Abb. 1) Nun stellt sich die Frage, ob sich eine solche Verteilung für die Simulation sozialer Netze eignet.

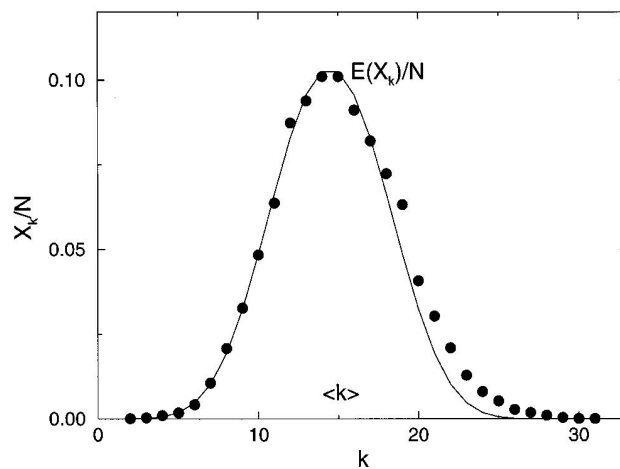


Abbildung 1: Simulation der Knotengradverteilung eines Erdős-Rényi-Zufallsgraphen.¹⁷ $N = 10000$, $p = 0.15\%$ (k = Knotengrad, N = Knotenzahl, X_k = Anzahl der Knoten mit Knotengrad k , p = Wahrscheinlichkeit das Knoten verbunden sind)

Die gleichmäßige Verteilung kommt dadurch zustande, dass es im Netzwerk genauso viele Knoten mit einem kleineren als auch mit einem größeren Knotengrad, bezogen auf den Durchschnitt, existieren. Diese Eigenschaft steht aber im Kontrast zu den Social Hubs, die bereits im Kapitel zwei vorgestellt wurden. Um diese ebenfalls in das Netzwerk zu integrieren, sollte eine Verteilung erzielt werden, bei der neben einer relativ gleichmäßig verteilten "Masse", auch eine Gruppe existiert, die einen deutlich höheren Knotengrad besitzt.

3.2.2 Generalisierte Zufallsgraphen

Eine Möglichkeit das Netzwerk besser an die Bedingungen sozialer Netzwerke anpassen zu können, besteht mit der Nutzung von generalisierten Zufallsgraphen (Generalized Random Networks). Bei dieser Art von Zufallsnetzwerken kann die Knotengradverteilung willkürlich gewählt werden. (Abb. 2)

¹⁷ vgl. Albert, Barabási, 2002, 57

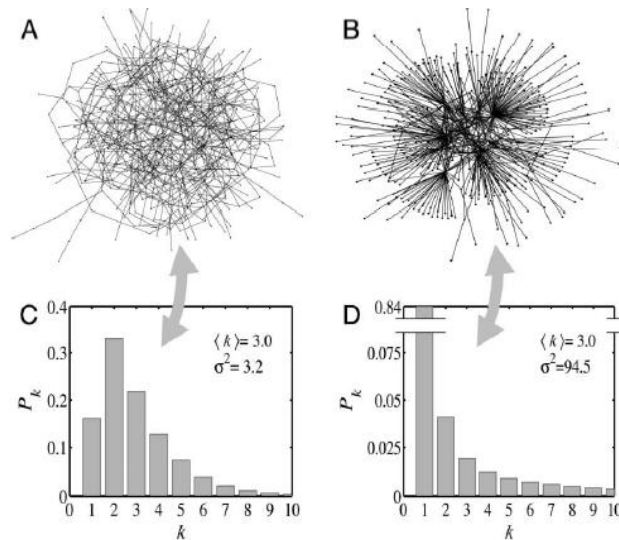


Abbildung 2: Zufallsgraphen mit festgelegter Knotengradverteilung.¹⁸ Durchschnittlicher Knotengrad 3. Niedrige Varianz (A) Hohe Varianz (B)

3.2.3 Power Law

Mit generalisierten Zufallsnetzwerken scheint es möglich ein Netzwerk zu erstellen, dass den Bedingungen von sozialen Netzwerken relativ nahe kommt. Doch eine wichtige Eigenschaft, die normale Zufallsgraphen nicht erfüllen, ist die der "Power Law" Verteilung¹⁹. Diese Verteilung wurde bei einer Vielzahl von sozialen Netzwerken nachgewiesen und findet auch in den Naturwissenschaften rege Verbreitung.²⁰ Die Verteilung zeichnet sich durch zwei besondere Eigenschaften aus, die auch für soziale Netzwerke wichtig sind.

Zum einen sind sie "skalenfrei", das bedeutet, dass die Verteilung des Knotengrades nicht von der Wahl der Skala abhängig ist. Es gibt also keinen charakteristischen Knotengrad, der sich auch bei einer Vergrößerung des Netzwerkes nicht ändert, sondern das Verhältnis der Knoten untereinander bleibt gleich. So bleibt auch die Struktur, ähnlich wie die Selbstähnlichkeit bei Fraktalen, immer gleich. (Abb. 3)

¹⁸ vgl. Watts, Dodds 2007, 8

¹⁹ "Power Laws" im Deutschen: Potenzgesetze der Skalengesetze

²⁰ vgl. Barabási 2003, 65-78

Zum anderen ist die Anzahl der Knoten mit einem großen Knoten-grad deutlich höher als bei normalen Zufallsgraphen. Ihre Anzahl nimmt nicht exponentiell ab. Auch dieses Merkmal deckt sich mit der Beobachtung von Social Hubs.



Abbildung 3: Wachstum einer Kochkurve.²¹ Dieses Fraktal wächst skalenfrei und bleibt sich selbst immer ähnlich.

3.2.4 Preferential Attachment

Eine relativ einfache Methode ein solches Netzwerk zu erzeugen ist das Barabási-Albert Modell, bei dem die Methode des "Preferential Attachment" benutzt wird. Bei dieser Methode werden die Knoten einzeln an das Netzwerk angefügt. Dabei ist die Wahrscheinlichkeit, sich mit einem bereits vorhandenen Knoten zu verbinden, abhängig von dessen Knotengrad. Je höher der Knotengrad ist, desto größer ist die Wahrscheinlichkeit, dass sich der neue Knoten mit ihm verbindet.

Das entstehende Netzwerk erfüllt die Bedingungen der Skalenfreiheit und der "Power Law" Verteilung. Im Vergleich mit zufälligen Netzwerken ergibt sich noch eine weitere Eigenschaft. Die durchschnittliche Pfadlänge²² ist deutlich kürzer. Das ist eine Eigenschaft, die perfekt auf eine Eigenart sozialer Netze passt, die bis jetzt noch nicht diskutiert wurde, das "Small World" Phänomen. Stanley Milgram zeigte bereits 1967 in einem Experiment, dass jeder Mensch im Schnitt über 6 andere Kontakte mit einem beliebig weit entfernten Menschen verbunden ist.²³ Auch wenn die

²¹ <http://de.wikipedia.org/wiki/Koch-Kurve>, Abgerufen am 24.07.2010

²² Die Durchschnittliche Pfadlänge bezeichnet den Durchschnitt der kürzesten Wege, die Knoten miteinander verbinden können. Im einfachsten Fall mit nur zwei Knoten wäre sie demzufolge eins.

²³ vgl. Albert, Barabási, 2002, 27ff

Pfadlänge von 6 umstritten ist, bleibt dennoch der Fakt, dass soziale Netze sich durch eine relativ kurze durchschnittliche Pfadlänge auszeichnen. Gerade in der elektronischen Kommunikation werden viele Hürden, die das "Small World" Phänomen normalerweise behindern²⁴, abgeschwächt. So konnten Jure Leskovec und Eric Horvitz 2008 eine durchschnittliche Pfadlänge von 6,6 im Netzwerk des Microsoft Messengers nachweisen.²⁵

3.2.5 Umsetzung des Netzwerkes in Netlogo

Nach der ausführlichen Beschäftigung mit der Theorie und der Auswahl eines Netzwerktypes, wird die Umsetzung dessen in Netlogo realisiert. Im Folgenden wird der dafür verwendete Algorithmus erörtert.²⁶ Dabei wird er nicht anhand der Programmierung, sondern systematisch an einem Erstellungsdurchgang in Netlogo beschrieben.

Als Ausgangspunkt dienen zwei Knoten, die mit einem Link miteinander verbunden sind. Danach wird bestimmt an welchen Knoten sich ein neu erstellter Knoten anbindet. Am Anfang dieser Phase wird eine Zufallszahl zwischen 0 und 1 erzeugt. Als Beispiel wählen wir jetzt 0,75. Danach wird ein Knoten gewählt und sein eigener Knotengrad (k_i) durch die Summe aller Knotengrade ($\sum_j k_j$) dividiert. In der Ausgangslage ergibt sich daher:

$$p = \frac{k_i}{\sum_j k_j} = \frac{1}{1+1} = \frac{1}{2}$$

Nun wird der berechnete Wert p mit der Zufallszahl verglichen. Ist der Wert von p größer, verbindet sich der neue Knoten mit dem ausgewählten Knoten, ist er kleiner schreitet die Berechnung fort.

²⁴ Dazu zählen beispielweise ethnische oder gesellschaftliche Barrieren.

²⁵ vgl. Horvitz, Leskovec 2008

²⁶ Der Algorithmus basiert auf einem Modul, was bereits in der Programmbibliothek von Netlogo intrigiert ist. vgl. Wilensky 2005

In unserem Fall ist die Zufallszahl größer als p . Demnach muss der nächste Schritt erfolgen. Dabei werden die bereits berechneten Knotengrade zu dem Knotengrad des nächsten ausgewählten Knotens addiert. Daher ergibt sich im zweiten Schritt:

$$p = \frac{\sum_i k_i}{\sum_j k_j} = \frac{1+1}{1+1} = 1$$

Jetzt ist der berechnete Wert p größer als die Zufallszahl. Der neu erstellte Knoten wird an den ausgewählten Knoten angebunden und die Abfolge beginnt von neuem.

Die Wahrscheinlichkeit, dass ein Knoten als Verbindungspartner ausgewählt wird, steigt mit seinem Knotengrad.

$$p_{\text{Auswahl}} = \frac{k_i}{\sum_j k_j}$$

In unserem Beispiel haben die beiden Ausgangsknoten durch ihren gleichen Knotengrad eine Wahrscheinlichkeit (p_{Auswahl}) von 50%. Nachdem der dritte Knoten hinzugefügt wurde, hat der Knoten mit den meisten Verbindungen immer noch eine Wahrscheinlichkeit von 50%, dass sich weitere Knoten mit ihm verbinden. Die beiden anderen Knoten besitzen nur noch eine Wahrscheinlichkeit von jeweils 25%. So entstehen die Social Hubs und eine große Menge "normaler" Knoten mit geringem Knotengrad.

3.3 Ein Netzwerk voller Dreiecke

Nachdem einem in der Mathematik ausführlich beschriebenen Weg bis hier gefolgt wurde, kann jetzt damit begonnen werden das Netzwerk weiter an die Bedingungen einer Simulation sozialer Netze anzupassen. Eine dieser Bedingungen ist es Cluster zu erstellen. Cluster bezeichnen einen Bereich des Netzwerks, der im

Vergleich mit homogenen Netzwerken viele Knotenverbindungen besitzt, die im Kreis verlaufen. Zusätzlich führen relativ wenige Verbindungen aus dem Cluster heraus. Solche Cluster entstehen in dem gewählten Modell des Preferential Attachment bereits rein zufällig. Messbar sind die Cluster durch den durchschnittlichen Clusterkoeffizienten (\bar{C}), dieser ist der Mittelwert aller lokaler Clusterkoeffizienten (C_i). Bestimmt wird dieser durch den Quotienten der Kanten die die Nachbarn eines Knotens verbinden (n) und der Anzahl der Kanten die theoretisch zwischen ihnen verlaufen könnten ($\frac{1}{2}k_i(k_i-1)$).

$$C_i = \frac{2n}{k_i(k_i-1)}$$

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i$$

Soziale Netzwerke zeichnen sich durch einen relativ großen Clusterkoeffizienten aus²⁷. Eine Bedingung die ansatzweise bereits durch die Methode des Preferential Attachment erfüllt wird, denn diese Netzwerke haben einen größeren Clusterkoeffizienten als rein zufällige Netzwerke. Allerdings ist der Koeffizient abhängig von der Netzwerkgröße. Bei sozialen Netzen hingegen bleibt er bei jeder Größe konstant.

Eine Möglichkeit auf die Clusterung Einfluss zu nehmen ist, bereits bei der Entstehung des Netzwerkes Cluster anzufügen. In dem Simulationsnetzwerk wird statt eines Knotens, ein Verbund von drei Knoten in jedem Schritt angefügt. Ein solches Dreieck ist bereits ein kleiner Cluster und jeder Knoten des Dreiecks wird einzeln mit dem Preferential Attachment Algorithmus an das Netzwerk angebunden. (Abb. 4 bis Abb. 6)

²⁷ Dieser Fakt, der mehrfach nachgemessen wurde, fußt auf einer Entdeckung von Mark Granovetter. Er bestimmte bereits 1973 unterschiedlich starke Verbindungen unter den Menschen. Die Theorie der "Strong and Weak Ties", findet bis heute Verwendung.

Netzwerkerstellung

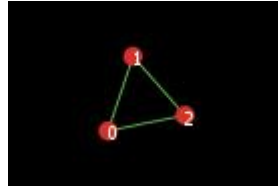


Abbildung 4: Ausgangslage der Netzwerkerstellung.²⁸ Das Dreieck $\{0,1,2\}$ bildet den ersten Cluster.



Abbildung 5: Zweiter Schritt der Netzwerkerstellung. Das Dreieck $\{3,4,5\}$ kommt neu hinzu.

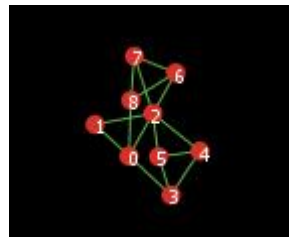


Abbildung 6: Dritter Schritt der Netzwerkerstellung. Das Dreieck $\{6,7,8\}$ kommt neu hinzu. Man erkennt bereits, dass sich aus Knoten $\{2\}$ ein Social Hub entwickelt. Sein Knotengrad ist bereits 6, während die meisten anderen Knoten nur einen Knotengrad von 3 besitzen.

Neben mathematischen Gründen für die Clusterbildung gibt es einen weiteren wichtigen Grund Dreiecke in das Netzwerk einzubinden. In sozialen Netzen existiert das Konzept der "Triadic Closure".²⁹ Dieses Konzept, was bereits vor 100 Jahren von dem deutschen Sozialwissenschaftler Georg Simmel entwickelt wurde,

²⁸ Bei den in der Arbeit verwendeten Abbildungen handelt es sich, sofern nicht anders angegeben, um eigene Ergebnisse von Simulationen und Netzwerkerstellungen mit Netlogo.

²⁹ sinngemäß im Deutschen: Dreiecksförmige Abschließung

gerät heute wieder in den Fokus.³⁰ "Triadic Closure" beschreibt den Umstand, dass die Dreiecksstruktur in sozialen Netzen sehr häufig auftaucht. Dies erscheint nachvollziehbar, wenn die Ausgangslage aus Abbildung 4 betrachtet wird. Ist ein Knoten $\{0\}$ mit dem Knoten $\{1\}$ und dem Knoten $\{2\}$ verbunden, ist es wahrscheinlich, dass die Knoten $\{1\}$ und $\{2\}$ ebenfalls verbunden sind.

3.4 Mega Hubs

In der Unterteilung der Begriffe von Michaela Klinger tauchen neben den Social Hubs und Experten Hubs auch Mega Hubs auf.³¹ Diese Hubs zeichnen sich durch eine weit überdurchschnittliche soziale Anbindung oder informellen Einfluss aus. Solche Mega Hubs entstehen aber nicht bei der Methode des Preferential Attachment. (Abb. 7)

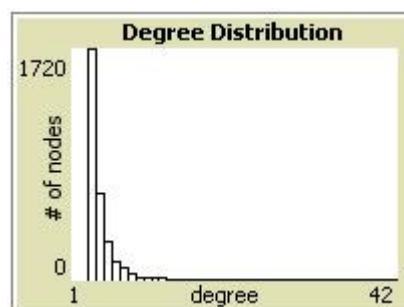


Abbildung 7: Knotengradverteilung der aktuellen Netzwerkerstellung bei 3000 Knoten. (Die Werte stellen eine einzelne Erstellung dar und keine Durchschnittswerte. Erneute Erstellungen können demzufolge abweichen)

Um die Erstellung zu ermöglichen, wird bei der Netzwerkentwicklung Einfluss auf die Zufallszahl genommen. Es wird ein Wert (a_{Abweich}) eingeführt, der bei jedem Durchgang auf die Zufallszahl addiert wird. Wenn a_{Abweich} negativ ist, steigt die

³⁰ vgl. Watts 2003, 60f

³¹ vgl. Klinger 2006, 86

Wahrscheinlichkeit, dass neue Knoten sich an den bereits existierenden, größten Social Hub anbinden. (Abb. 8)

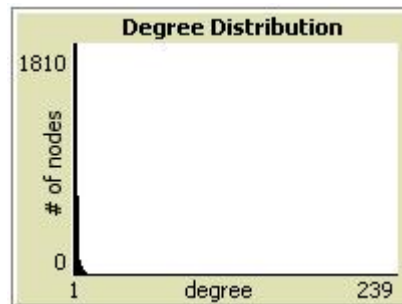


Abbildung 8: Knotengradverteilung der aktuellen Netzwerkerstellung bei 3000 Knoten und einem Abweichwert von -0,05. Man erkennt, dass der maximale Knotengrad von 41 auf 238 angewachsen ist. (Das Ausgabefenster ist immer um eins größer als der maximale Knotengrad.)

Ist a_{Abweich} positiv, steigt die Wahrscheinlichkeit, dass die Zufallszahl auch nach der Aufsummierung aller Knoten nicht erreicht wird. In diesem Fall wird der Knoten an einen rein zufällig ausgewählten Knoten angebunden. So ergibt sich die gegenteilige Verteilung. Anstatt das Wachsen eines Mega Hubs zu beschleunigen, wird das Netzwerk gleichverteilter. (Abb. 9)

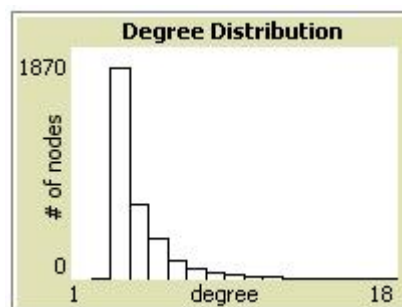


Abbildung 9: Knotengradverteilung der aktuellen Netzwerkerstellung bei 3000 Knoten und einem Abweichwert von 0,05. Der maximale Knotengrad hat sich von 41 auf 17 verringert.

Da die Einführung von a_{Abweich} eine deutliche Abweichung von der mathematischen Theorie darstellt und die Simulation stark beeinflussen kann, ist diese Funktion in der Netzwerkerstellung optional.

3.5 Dimensionierung des Knotengrades

Nachdem die Struktur des Netzwerkes festgelegt ist und modifizierbar gestaltet wurde, wird sich mit der Dimensionierung des Knotengrades befasst. Um das Netzwerk bestimmen zu können, wird der Frage nachgegangen, wie viele Verbindungen die Knoten besitzen sollen. Eine Möglichkeit wäre es das Netzwerk so zu erstellen, dass die Verbindungen den Kontakten entsprechen, die Menschen in sozialen Netzwerken besitzen.

3.5.1 Soziale Kontakte

Zu den sozialen Kontakten unterschiedlicher Kommunikationstypen hat Malcolm Gladwell in seinem Buch "Tipping Point" ein interessanten Versuch beschrieben.³² Er wählte aus einem Telefonbuch rein zufällig 250 Namen aus. Danach überprüfte er bei allen Namen, ob er jemanden mit gleichem Namen kennt. Dabei betrachtete er "kennen" sehr locker. Das bedeutet, dass er jede Form der Bekanntschaft, ob gute Freunde oder entfernte Bekanntschaften, dazu zählte. Diesen Test lies er auch weitere 400 Probanden unterschiedlicher Alters-, Berufs-, und Sozialschichten durchführen. 25 der Probanden kannten weniger als 20 Leute, und 12 kannten mehr als 90 Leute. Das bedeutet, dass über 90% der Probanden zwischen 20 und 90 Leute mit dem gleichen Namen auf der Liste kannten. Wenn man bedenkt, dass die Liste rein zufällig erstellt wurde, ist diese Zahl bereits sehr groß. Die Gesamtzahl aller Kontakte müsste diese Zahlen noch deutlich übersteigen. Für das Simulationsnetzwerk bedeutet dies, dass jeder Knoten wenigstens einige Dutzend Verbindungen besitzen müsste. Im Fall von Social Hubs könnten sie sogar einige Hundert

³² vgl. Gladwell 2002, 52 - 56

Verbindungen besitzen. Diese enorme Anzahl an Verbindungen würde eine starke Anpassung des Netzwerkes verlangen. Im aktuellen Netzwerkentwurf haben die kleinsten Knoten nur drei Verbindungen. Somit müssten an jeden Knoten noch dutzende Verbindungen angefügt werden. Gegen diese Anpassung sprechen aber zwei wichtige Gründe. Zum einen liegen keine genauen Zahlen vor wie viele Verbindungen jeder Knoten bekommen sollte und zum anderen wird auch die Performance der Simulationslösung durch die enorme Steigerung der Verbindungen deutlich schlechter.

3.5.2 Weitergabennetzwerk

Nach den bisherigen Vorbetrachtungen ist es sinnvoll nach einem anderen Weg zu suchen. Dabei helfen die Forschungen und Erhebungen im Bereich des viralen Marketings und des Word of Mouth. Bei diesen Marketingstrategien ist es wichtig zu ermitteln, wer wie oft die Botschaften durchschnittlich weitergibt. (Abb. 10) Mit diesen Zahlen kann das Netzwerk so angepasst werden, dass die Verbindungen nicht die Bekanntschaften der Knoten untereinander darstellen, sondern die durchschnittlichen Weitergabepartner im Internet. Das Netzwerk wandelt sich somit von einem kompletten sozialen Netzwerk zu einem Weitergabennetzwerk, bei dem Knoten und Verbindung über die keine Informationen ausgetauscht werden, ausgeblendet werden.

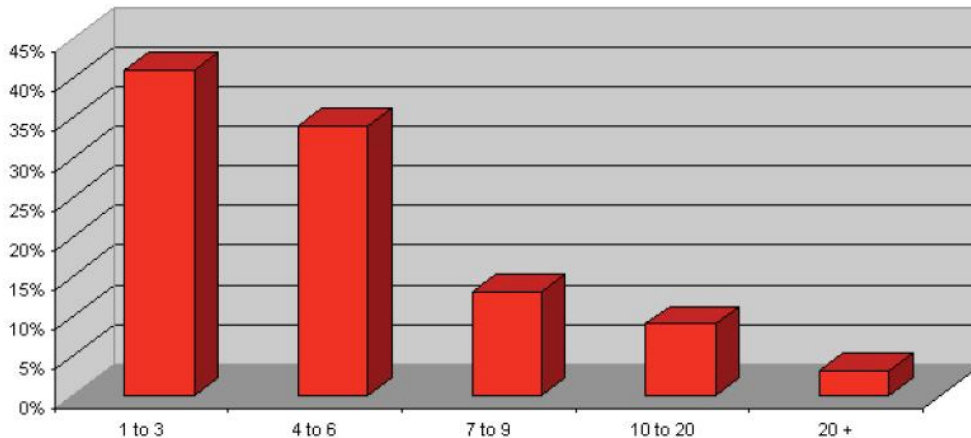


Abbildung 10: Anzahl der Leute mit denen Internet Nutzer Inhalte teilen.³³

Es muss aber davon ausgegangen werden, dass die verbundenen Knoten in beide Richtungen die gleich hohe Motivation besitzen, Inhalte zu teilen. Beispielsweise bei zwei Knoten A und B, die jeweils ihre Inhalte durchschnittlich drei mal teilen. Wenn einer der drei Weitergabepartner von Knoten A, Knoten B ist, muss umgekehrt auch einer der Weitergabepartner von Knoten B, Knoten A sein. Diese Annahme ist vertretbar, da es durchaus wahrscheinlich ist, dass bei einer Informationsweitergabe der Weitergabepartner bewusst ausgewählt wird und eben dieser Partner die gleiche Überlegung anstellt. So teilen beispielsweise zwei Autofanater die Informationen zu Autos bevorzugt untereinander und nicht rein willkürlich im restlichen Bekanntenkreis. In diesem Fall würde der restliche Bekanntenkreis in der Simulationslösung keine Beachtung mehr finden.

3.5.3 Die Erstellung von Social Hubs

Im aktuellen Netzwerkentwurf weichen die Knotengrade noch stark von den statistisch erhobenen Werten ab. (Abb. 11)

³³ vgl. Viral Manager - How Users Share Viral Content o.J., 2

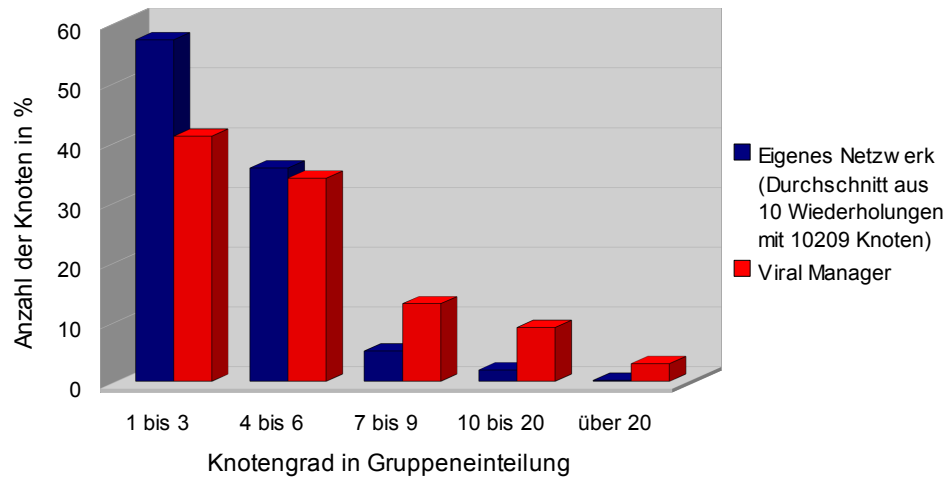


Abbildung 11: Vergleich der Knotengrade des Simulationsnetzwerkes mit den Daten aus Abbildung 10.

Die größten Abweichungen lassen sich in der ersten Gruppe und bei Knotengraden über 7 feststellen. Um diese Abweichungen anzugleichen, wird eine neue Funktion (CSH - Create Social Hubs) in die Simulationslösung eingeführt. Diese Funktion erzeugt durch neue Verbindungen zwischen den Knoten weitere Social Hubs unterschiedlicher Knotengrade. Dabei werden Knoten aus der ersten Gruppe mit anderen Knoten aus höheren Gruppen verbunden, um eine vorgegebene prozentuale Verteilung zu erreichen. Da die Methode aber nur näherungsweise arbeitet und die erste Gruppe durch die Erhöhung der Knotengrade stetig abnimmt, kann keine beliebige Verteilung bestimmt werden. Zur Angleichung an die statistisch erhobenen Werte genügt sie jedoch völlig. (Abb. 12)

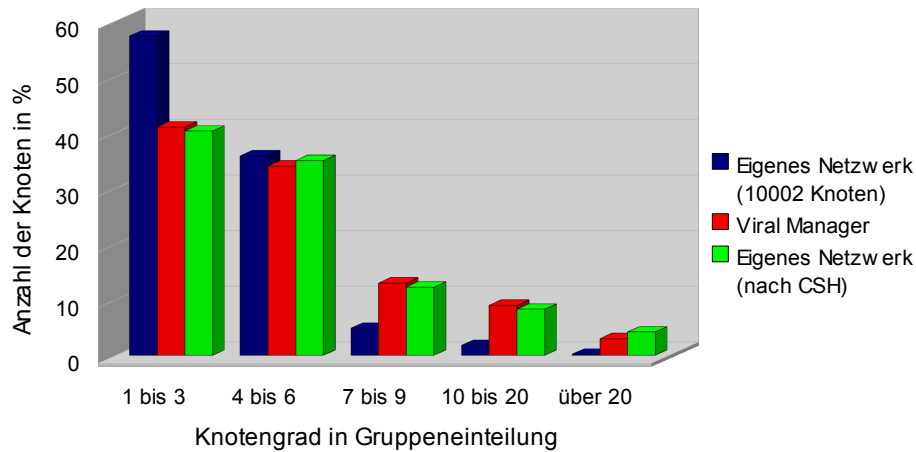


Abbildung 12: Erneuter Vergleich mit den Daten aus Abbildung 10. Bei CSH sollte Gruppe 3 und 4 jeweils um 8% und Gruppe 5 um 2% angehoben werden.

Die Funktion CSH greift ähnlich stark wie die Funktion zum Erstellen von Mega Hubs in die Knotengradverteilung ein. (Abb. 13) Dadurch entstehen einige neue Eigenarten der Verteilung. Zum einen hebt die Funktion CSH den Knotengrad einzelner Knoten nicht beliebig hoch an. Bei sehr großen Netzwerken kann also eine Lücke zwischen den angehobenen Knoten und den durch "Preferential Attachment" erzeugten Social Hubs entstehen. Zum anderen folgt die Verteilung im beeinflussten Knotengradbereich nun nicht mehr der Power Law Verteilung. (Abb. 14) Daher ist auch die Funktion CSH optional.

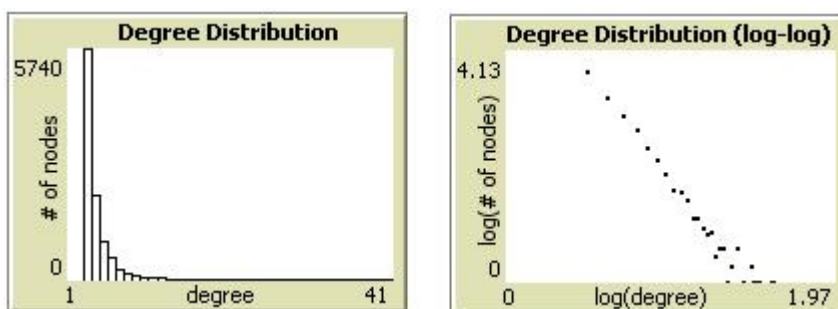


Abbildung 13: Kontengradverteilung vor CSH. (10002 Knoten) In der logarithmischen Darstellung ist die Gerade der Power Law Verteilung gut erkennbar.

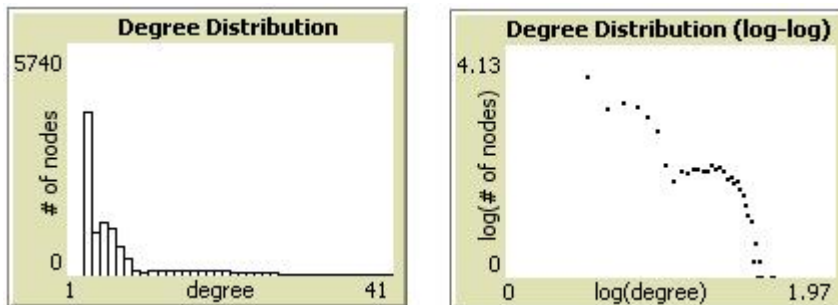


Abbildung 14: Knotengradverteilung nach CSH. (10002 Knoten) Man erkennt, dass die Knoten nicht bis zum maximalen Knotengrad angehoben werden (links) und die Abweichungen von der Power Law Verteilung (rechts).

3.6 Obergrenze der Social Hubs

Wie bereits im Kapitel 3.2.3 diskutiert wurde, wachsen die Knotengrade mit der Gesamtgröße des Netzwerkes mit. Bei größeren Netzwerken von einigen zehntausend oder sogar hunderttausend Knoten kann der größte Knotengrad durchaus bei weit über hundert Verbindungen liegen.³⁴ Gerade wenn die Methode zur Erzeugung von Mega Hubs aus Kapitel 3.4 verwendet wird, kann der maximal Knotengrad bei einigen hundert, sogar tausenden Verbindungen liegen. Da das Simulationsnetzwerk als Weitergabennetzwerk betrachtet wird, sollte sich kritisch mit sehr großen Knotengraden auseinander gesetzt werden.

In einem sozialen Netzwerk bedeutet ein hoher Knotengrad eine große Anzahl sozialer Kontakte. In unserem Weitergabennetzwerk ist er allerdings gleichbedeutend mit einer großen Zahl an Weitergaben. Eine Zahl von über hundert Weitergaben bedeutet aber einen enormen Zeit und Arbeitsaufwand, der ab einem gewissen Punkt nicht mehr vertretbar erscheint. Solch ein Arbeitsaufwand wird meist nur von aktiv Werbetreibenden

³⁴ Der maximale Knotengrad ist jedoch vom Zufall abhängig. Bei einer Erstellung eines Netzwerkes mit 20.088 Knoten lag der größte Knotengrad bereits bei 75, bei einem weiteren Versuch erreichte er diese Größe erst bei über 140.000 Knoten. Die Power Law Verteilung bleibt dennoch immer näherungsweise bestehen.

aufgewendet.³⁵ Daher scheint es sinnvoll den maximalen Knoten-grad nach oben hin zu begrenzen. Doch eine genaue Grenze ist kaum definierbar. Gerade im Internet, wo es immer mehr Möglichkeiten gibt Informationen einer großen Menge von unbekannten oder bekannten Menschen zugänglich zu machen. Von E-Mail über Blogs hin zu Twitter, wird es immer einfacher Informationen zu teilen. Eine Obergrenze ist daher stark abhängig von der Struktur beziehungsweise der Nutzung von "Massenkommunikationsmitteln" in der zu untersuchenden Zielgruppe. Daher kann die Obergrenze flexibel in der Simulationslösung eingestellt werden.

³⁵ Beispielsweise bei der Firma BzzAgents, die Menschen für Word of Mouth Marketing wirbt. Diese werben teilweise mehrere Stunden in der Woche und betreiben aktiv Mundpropaganda. Sie haben allerdings eine Motivation, die dem normalen Kommunikationsteilnehmer fehlt. (Walker 2004)

4 Simulationsmethoden

4.1 Verwendung des Netzwerks

Nachdem das Netzwerk erstellt wurde, wird die Verbreitung von Informationen darin untersucht und simuliert. Es werden die Fragen geklärt, wie die Knoten Informationen weitergeben sollen, wie sie sie annehmen sollen und was nach einer Annahme passiert. Dabei haben sich bei bisherigen Simulationen zwei Modelle bewährt, die für die Nutzung in Netlogo geeignet wären. Zum einen das Threshold-³⁶ und zum anderen das SIR-Modell.³⁷ Beide Modelle sollen im Folgenden vorgestellt und eine Umsetzung in Netlogo erarbeitet werden.

4.2 Binäre Entscheidungen

Das Threshold- und das SIR-Modell haben eine gemeinsame Vorbedingung. Es gibt für die Knoten nur eine Wahl zwischen zwei Alternativen, entweder Information übernehmen oder nicht übernehmen. Das bedeutet, dass der komplexe Entscheidungsprozess auf eine binäre Entscheidung reduziert wird. Binäre Entscheidungen reichen allerdings völlig aus, um eine große Zahl von Entscheidungssituationen zu simulieren.³⁸

Diese Einschränkung stellt für unser Simulationstool kein großes Hindernis dar. Die Verbreitung der Information, ob die Knoten die Informationen angenommen haben oder nicht, stellt den Betrachtungsschwerpunkt der Simulation dar. Wenn die Knoten die Information aufgenommen haben, können immer noch weitere Betrachtungen angestellt werden. Beispielsweise wie groß die Wirkung auf sie ist. Dazu ist kein komplexeres Entscheidungssystem notwendig.

Durch die Beschränkung auf eine binäre Entscheidung, ist eine Auswahl zwischen mehreren Möglichkeiten nicht mehr

³⁶ Threshold Deutsch: Schwelle

³⁷ SIR: Susceptible Infected Recovered Deutsch: Anfällig-Infiziert-Genesen

³⁸ vgl. Watts, Dodds 2007, 2

simulierbar. Beispielsweise die gleichzeitige Verbreitung von konkurrierenden Informationen.³⁹

4.3 Threshold Modell

Das Threshold Modell beschreibt einen Schwellenwert bei dem die Knoten eine Information übernehmen. Wenn die Anzahl der Nachbarn eines Knotens, die die Information besitzen, den Schwellenwert des Knotens übersteigt, übernimmt er die Information. Ein Knoten mit dem Schwellenwert von 0,5 übernimmt beispielsweise die Information, wenn mehr als die Hälfte der Nachbarn ebenfalls die Information übernommen haben.

4.3.1 Grundlagen des Threshold

Dieses recht einfach klingende Modell beruht nicht auf reiner Willkür von Wissenschaftlern, die nach einer Simulationslösung suchen, sondern hat eine psychologische Berechtigung. Zur Verdeutlichung soll das Konformitätsexperiment von Solomon Asch aus den fünfziger Jahren dienen.⁴⁰

Bei diesem recht simplen Experiment wurde versucht den Konformitätsdrang oder anders ausgedrückt den Gruppenzwang nachzuweisen. Den Probanden wurden Abbildungen von Linien vorgeführt. (Abb. 15)

³⁹ Zumindest im Rahmen dieser Arbeit. Mit einer starken Modifizierung der Simulationsmodelle wäre auch eine solche Simulation denkbar.

⁴⁰ vgl. o.V. <http://www.analytictech.com/mb109/slides/asch.pdf> Abgerufen am 24.08.2010

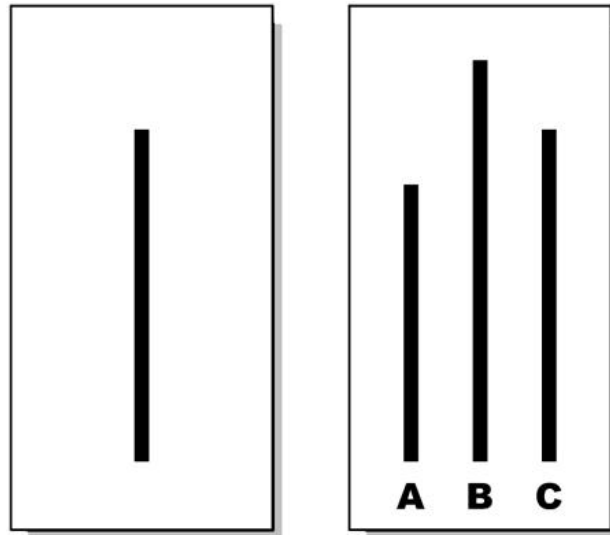


Abbildung 15: Versuchsbild des Asch-Experiments.⁴¹ Balken C entspricht dem Versuchsbalken. Oder doch nicht?

Die Aufgabe bestand nun darin die Linie zu finden, welche in der Länge der linken Testlinie entsprach. Die Besonderheit des Tests lag darin, dass immer nur ein echter Proband im Raum war und alle anderen Personen Vertraute des Versuchsleiters waren. Konnten alle Personen ehrlich ihre Meinung äußern, hatten die Probanden bei 18 Wiederholungen mit unterschiedlichen Linien nur eine Fehlerrate von unter einem Prozent. Sobald allerdings die Eingeweihten einheitlich eine falsche Antwort gaben, stieg die Fehlerrate bei den Versuchspersonen auf über 30%.⁴² Somit beugten sie sich in rund einem Drittel der Fälle dem Gruppendruck und machten offensichtliche Falschaussagen.

An diesem Experiment erkennt man die Existenzgrundlage des Threshold Modells. Menschen neigen dazu sich ab einem gewissen Schwellenwert der Meinung anderer anzuschließen. Dennoch bleibt die Betrachtung des Schwellenwertes an sich. Im Experiment von Asch fiel die Fehlerquote deutlich auf 5% ab, nachdem

⁴¹ http://upload.wikimedia.org/wikipedia/commons/4/47/Asch_experiment.png Abgerufen am 30.07.2010

⁴² vgl. o.V. http://www.psychologie.wiso.uni-erlangen.de/lehrveranstaltungen/ss06/sozialpsychologie/sozialereinfluss1_skript.pdf Abgerufen am 30.07.2010

nur einer der Mitwisser ebenfalls die richtige Antwort gab. Daran erkennt man, dass die Akzeptanz einer offensichtlichen Fehlinformation einen sehr hohen Schwellenwert besitzt.⁴³ Im Simulationsmodell muss allerdings nicht zwangsläufig von solch einem Extremfall ausgegangen werden. Daher kann man eher von Threshold Werten um den Wert 0,5 ausgehen.⁴⁴ Andere Simulationen von Informationsausbreitung wurden bereits erfolgreich mit einem Threshold Wert von 0,18 durchgeführt.⁴⁵

4.3.2 Umsetzung des Threshold Modells in Netlogo

Um die Umsetzung des Threshold Modelles in Netlogo zu erläutern, soll der Prozess an einem einfachen Beispiel verdeutlicht werden. Betrachten wir ein kleines Netzwerk in dem ein Beispielknoten {1} neben den Nachbarknoten {2} und {3} noch vier weitere Nachbarn besitzt. Insgesamt beträgt der Knotengrad des Knotens {1} also sechs.

Nun trägt der Knoten {2} die Information an den Beispielknoten {1} heran. Dieser vergleicht nun den Quotienten aus informationstragenden Nachbarn (k_{inf}) und der Gesamtanzahl der Nachbarn (k) mit einem vor der Simulation festgelegten Threshold (Φ). Ist der Wert größer als der Threshold übernimmt der Knoten die Information. In unserem Fall sei der Threshold 0,3.

$$\frac{k_{inf}}{k} = \frac{1}{6} = 0,1\bar{6} < 0,3 = \Phi$$

In dieser Lage übernimmt der Knoten {1} die Information nicht. Knoten {2} bleibt aber weiterhin Träger der Information, wenn auch nur passiv. Das bedeutet, dass er die Information nicht noch

⁴³ Bei n Gruppenteilnehmern liegt der Threshold (Φ) somit bei: $\Phi \geq \frac{n-1}{n}$

⁴⁴ vgl. Latané L ´ Herrou 1996

⁴⁵ vgl. Watts, Dodds 2007, 5

Dabei wurde allerdings ausschließlich mit dem Threshold Modell gearbeitet.

einmal an den Beispielknoten {1} herantragen wird, aber weiterhin zu den informationstragenden Nachbarn gezählt wird.

Im zweiten Schritt unseres Beispiels trägt auch der Knoten {3} die Information an den Beispielknoten {1} heran. Daher ergibt sich analog zum ersten Schritt:

$$\frac{k_{inf}}{k} = \frac{2}{6} = 0,3\bar{3} > 0,3 = \Phi$$

In diesem Fall übernimmt der Knoten {1} die Information da sein Threshold übertroffen wurde.

4.3.3 Wahl des Thresholds

Der Threshold ist vor der Simulation frei vom Benutzer zu wählen. Diese Wahl hängt hauptsächlich von der Art der Information ab, die verbreitet werden soll. Wie bereits das Asch-Experiment zeigte, gibt es einen sehr großen Threshold bei offensichtlichen Fehlinformationen. Ein deutlich niedrigerer Threshold ist zu erwarten, wenn die Information etwas triviales wie eine Produktinformation oder Nachrichtenmeldung ist.

Weitere Faktoren für die Auswahl des Thresholds sind Aufwand und Nutzen, der sich für den Nutzer ergibt. Beispielsweise bei der Verbreitung eines Tools wie ICQ⁴⁶, dessen Nutzen sehr von der Anzahl der Kontakte, die es ebenfalls nutzen, abhängt. In diesem Fall ist ein größerer Threshold zu erwarten, als bei einem Tool dessen Nutzen von den umgebenen Knoten unabhängig ist. Analog verhält es sich mit dem Aufwand den der Nutzer für die Information aufwenden muss. Muss er sich erst registrieren oder Zusatzsoftware installieren, ist ein höherer Threshold zu erwarten als bei direkt zugänglichen Informationen oder Applikationen.

Der letzte Faktor, der bei der Wahl des Thresholdes zu beachten ist, sind die persönlichen Eigenschaften der

⁴⁶ <http://www.icq.de/de.html> Abgerufen am 31.07.2010; ICQ ist ein Kommunikationstool bei dem man mit Freunden, die es ebenfalls nutzen, einfach in Verbindung treten kann.

Netzwerkteilnehmer. Es gibt für einen Informationstyp keinen absoluten Threshold, sondern nur eine charakteristische Verteilung um den Durchschnittswert, denn jede Person ist verschieden.⁴⁷ (Abb. 16) Somit sollte die Simulation mit unterschiedlichen Thresholdwerten wiederholt durchgeführt werden, um die Auswirkungen zu untersuchen.

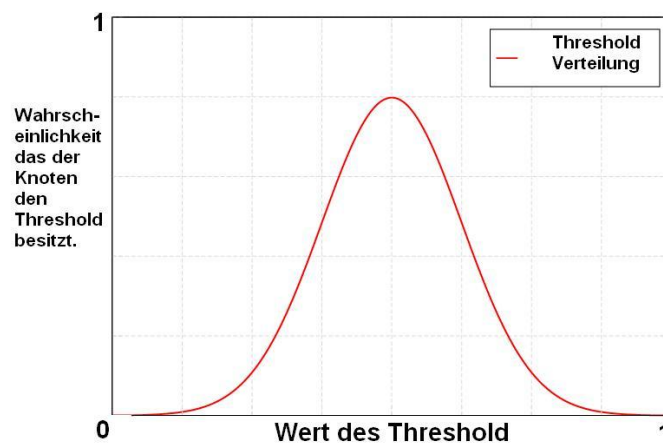


Abbildung 16: Eine mögliche Thresholdverteilung.⁴⁸

4.3.4 Simulationsentscheidende Thresholdwerte

Bedingt durch die Netzwerkstruktur, gibt es drei unterschiedliche Wertebereiche für den Threshold (Φ), die starken Einfluss auf das Simulationsergebnis haben. Das ist durch das vorherrschende Netzwerkelement, dem Dreieck und dem dominierenden Knoten-grad von drei bestimmt.

Der erste Bereich liegt bei:

$$0 \leq \Phi \leq \frac{1}{3}$$

⁴⁷ vgl. Watts 2003, 227

⁴⁸ In Anlehnung an: <http://upload.wikimedia.org/wikipedia/de/5/50/Dichtefunktion.png> Abgerufen am 31.07.2010. Nach Watts 2003, 227

In diesem Fall reicht bereits ein informationstragender Nachbar-knoten aus, damit die Dreiecke ebenfalls die Information übernehmen. (Abb. 17)

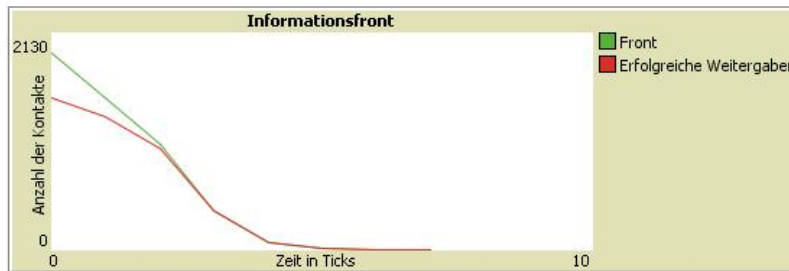


Abbildung 17. Informationsausbreitung mit einem Threshold von 0,33. Simuliert wurde mit 5001 Knoten und 15% zufällig verteilten Informationsträgern. Mit der Front werden die Nachbarn aller informationstragender Knoten dargestellt. Die erfolgreichen Weitergaben stellen alle Weitergaben an diese Nachbarn dar.

In der Simulation ist zu erkennen, dass zu Beginn fast alle Nachbarn durch den Threshold ebenfalls zu Informationsträgern wurden. Lediglich bei Social Hubs wurde der Threshold nicht erreicht.⁴⁹ Aber mit zunehmender Verbreitung der Information wurde zunehmend auch der Threshold von Social Hubs erreicht, so dass ab einem gewissen Punkt jeder mögliche Nachbar auch infiziert wurde.⁵⁰ Am Ende der Simulation war jeder Knoten im Netzwerk Träger der Information.

Der zweite Bereich liegt bei:

$$\frac{1}{3} < \Phi \leq \frac{2}{3}$$

In diesem Bereich müssen in den Dreiecken bereits zwei von drei Nachbarn Informationsträger sein, um einen weiteren Knoten zu infizieren. Dadurch weichen die Simulationsergebnisse deutlich vom ersten Bereich ab. (Abb. 18 und Abb. 19)

⁴⁹ Bei einem Threshold der sich 0 annähert, wurden auch die Social Hubs bereits zu Beginn zu Informationsträgern.

⁵⁰ "Infizieren" wird im folgenden analog zu Information übertragen verwendet. Diese Bezeichnung ist bei Virusausbreitung in Netzwerken üblich.

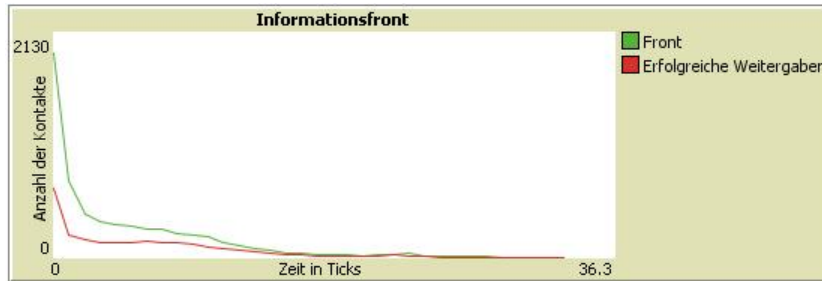


Abbildung 18: Informationsausbreitung mit einem Threshold von 0,34. Simuliert wurde mit 5001 Knoten und 15% zufällig verteilten Informationsträgern.



Abbildung 19: Informationsausbreitung mit einem Threshold von 0,66. Simuliert wurde mit 5001 Knoten und 15% zufällig verteilten Informationsträgern.

Der Unterschied zum ersten Bereich ist offensichtlich. Beim Ändern des Thresholds von 0,33 auf 0,34 verlängerte sich die Ausbreitungszeit um mehr als das Dreifache und am Ende der Simulation waren nur 66% der Knoten Informationsträger. Bei einem Threshold von 0,66 endet die Informationsausbreitung deutlich früher. Am Ende der Simulation sind nicht einmal 20% der Knoten Informationsträger. Abzüglich der anfänglichen 15% wurden somit knapp 5% neu infiziert.

Der dritte Bereich liegt bei:

$$\frac{2}{3} < \Phi \leq 1$$

In diesem Bereich müssen alle Nachbarn der häufigsten Knoten mit Knotengrad drei bereits Informationsträger sein. (Abb. 20)



Abbildung 20: Informationsausbreitung mit einem Threshold von 0,67. Simuliert wurde mit 5001 Knoten und 15% zufällig verteilten Informationsträgern.

Der Trend, der sich bereits mit einem Threshold von 0,66 abgezeichnet hat, verstärkt sich mit einem Wert von 0,67 noch deutlich. Es werden gerade mal 0,4% der Knoten neu infiziert und die Ausbreitung endet bereits kurz nach ihrem Start. Bei einer Simulation mit einem noch größeren Threshold sind keine anderen Ergebnisse zu erwarten.

Die Sprünge in den Ergebnissen bei der Veränderung des Thresholdes von 0,33 auf 0,34 und 0,66 auf 0,67 müssen bei der Wahl des Thresholdes bedacht werden. Andernfalls wird die Netzwerkstruktur bei der Simulation vernachlässigt und die Ergebnisse der Simulation könnten fehlinterpretiert werden.

4.4 Die SI-, SIS- und SIR-Modelle

Die SI-, SIS- und SIR- Modelle entstammen der mathematischen Modellierung der Epidemiologie. Dabei wird versucht die Ausbreitung von Infektionskrankheiten in Netzwerken zu simulieren. Die Ansätze können auch für die Simulation von Informationsausbreitung benutzt werden, wenn auch nicht jedes Modell passend ist. Daher wird zunächst die beste Variante ausgewählt.

4.4.1 Auswahl des Modells für die Simulation

Das einfachste der Modelle ist das SI-Modell. Dabei gibt es zwei Status, die die Knoten einnehmen können. Zum einen "S" für Susceptible⁵¹ und zum anderen "I" für Infectious⁵² oder Infected.⁵³ In der Ausgangslage besteht das Netzwerk aus anfälligen Knoten und einer festgelegten Anzahl von infektiösen Knoten. Nach dem Kontakt mit einem infektiösen Knoten wird ein anfälliger Knoten ebenfalls infiziert. Da keinerlei Möglichkeit besteht zu gesunden, wird früher oder später das gesamte Netzwerk infiziert. Diese einfache Variante bietet sich daher nicht für eine realistische Simulation von Informationsausbreitung an.

Das SIS-Modell ergänzt die Abfolge der Ausbreitung um einen weiteren Schritt. Hier werden die Knoten nach ihrem infizierten Status wieder in den anfälligen Status versetzt. Somit ist es möglich, dass die Krankheit das komplette Netzwerk infiziert, dass sie stabil im Netzwerk verbleibt ohne jedoch alle Knoten zu infizieren oder dass die Krankheit komplett verschwindet. Allerdings ist der Umstand, dass die Knoten nach ihrer Infizierung wieder anfällig werden, für eine Simulation von Informationsausbreitung unrealistisch. Es ist nicht nachvollziehbar, dass ein Kommunikationsteilnehmer, nachdem er eine Information besessen und weitergegeben hat, plötzlich alles wieder vergisst und in der gleichen Situation wieder gleich verfährt.

Im SIR-Modell⁵⁴ wird ein neuer Status eingeführt. Die Knoten können nach ihrer Infizierung resistent werden. Das bedeutet, dass sie nicht wieder infiziert werden können und sie den Virus auch nicht erneut weitergeben. Nun hängt es davon ab wie infektiös das Virus ist und mit welcher Wahrscheinlichkeit die Knoten resistent werden können, ob das Netzwerk teilweise, komplett oder überhaupt nicht infiziert wird. Wenn man nun annimmt, dass die Knoten nach einer Infizierung sicher resistent werden, scheint dieses Modell geeignet um die Informationsausbreitung zu

⁵¹ Deutsch: anfällig

⁵² Deutsch: infektiös

⁵³ Deutsch: infiziert

⁵⁴ Dieses Modell geht auf Kermack, McKendrick 1927 zurück.

simulieren. Die Kommunikationsteilnehmer würden sich, nachdem sie einmal die Information angenommen und weitergegeben haben, resistent gegenüber der weiteren Übertragung verhalten.

4.4.2 Umsetzung des SIR-Modells in Netlogo

Die Umsetzung in Netlogo wird wieder an einem relativ simplen Beispiel verdeutlicht. Betrachten wir ein Netzwerk mit einem Beispielknoten {1}, der mit zwei Nachbarknoten {2} und {3} verbunden ist. In der Ausgangslage sei der Knoten {2} wieder Träger der Information und gibt sie an den Knoten {1} weiter. Dieser nimmt nun die Information mit einer Wahrscheinlichkeit (p_A) an, die vom Ersteller der Simulation festgelegt werden kann. Danach wird der Knoten {2} resistent. Das bedeutet, dass er weder erneut Träger der Information werden kann, noch dass er die Information noch einmal an seine Nachbarn übergeben wird. Nun gibt der Knoten {1} mit einer bestimmten, ebenfalls vom Ersteller der Simulation festgelegten Weitergabewahrscheinlichkeit (p_W), die Information an den Knoten {3} weiter. Daraufhin wird der Knoten {1} ebenfalls resistent. Am Ende der Simulation kann mit der Anzahl der resistenten Knoten die Reichweite der Informationsverbreitung bestimmt werden. (Abb. 21)

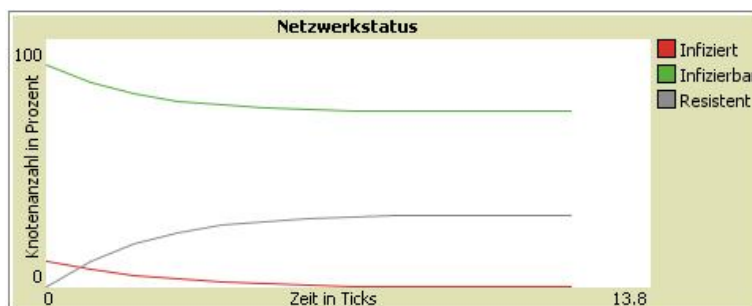


Abbildung 21: Beispiel einer SIR-Simulation. Simuliert wurde mit 5001 Knoten, $p_A = 0,2$, $p_W = 1$ und 10% zufällig gewählten Informationsträgern. Das Endergebnis sind in diesem Versuch 28,8% resistente Knoten.

4.5 Kombination der Modelle

An diesem Punkt soll mit dem Simulationstool einen neuer Weg beschritten werden. Anstatt eine der beiden vorgestellten Simulationsmethoden einzeln durchzuführen, werden diese miteinander kombiniert. Dies erscheint logisch, da bei einer Verwendung des SIR-Modells die Grundlage des Threshold, der Gruppenzwang, nicht außer Acht gelassen werden kann. Im umgekehrten Fall, einer Verwendung des Threshold, kann auch das SIR-Modell nicht außer Acht gelassen werden. Denn damit kann im Gegensatz zum Threshold auch die Überzeugungskraft und der Einfluss einzelner Knoten simuliert werden. Am deutlichsten werden die Defizite der Modelle, die mit einer Kombination ausgeglichen werden sollen, bei der Infizierung der Social Hubs. Im Threshold Modell sind sie nur schwer zu infizieren, da sie aufgrund ihres hohen Knotengrades den Threshold nur selten erreichen. Im SIR-Modell hingegen sind sie sehr leicht zu infizieren, da sie aufgrund ihrer vielen Verbindungen eine große Zahl an Infizierungsmöglichkeiten besitzen.

Für die Kombination der Modelle in Netlogo ist die Ausführungsreihenfolge entscheidend. Zuerst wird überprüft, ob die Weitergabewahrscheinlichkeit erreicht wird, wenn nicht, kann es weder eine Weitergabe via Threshold noch eine Weitergabe durch das SIR-Modell geben. Danach wird die Weitergabe durch den Threshold durchgeführt. Erst wenn diese scheitert, wird die Weitergabe mit dem SIR-Modell durchgeführt.

Diese Reihenfolge erklärt sich aus dem logischen Zusammenhang, dass eine Prüfung der Annahmewahrscheinlichkeit nicht nötig ist, wenn der Gruppenzwang den Knoten sowieso zur Übernahme bewegt. Ein weiterer Grund ist, dass sich die Weitergaben, die es ohne den Threshold gegeben hätte, besser abschätzen lassen als anders herum. Denn das SIR-Modell ist an einfache Wahrscheinlichkeiten gebunden, während der Threshold stark von der aktuellen strukturellen Verteilung der Information im Netzwerk abhängt. Um die Auswirkung der Kombination der Simulationsmodelle schnell und einfach simulieren zu können, ist der Threshold in der Simulation einfach ab- und zuschaltbar. (Abb. 22 und Abb. 23)

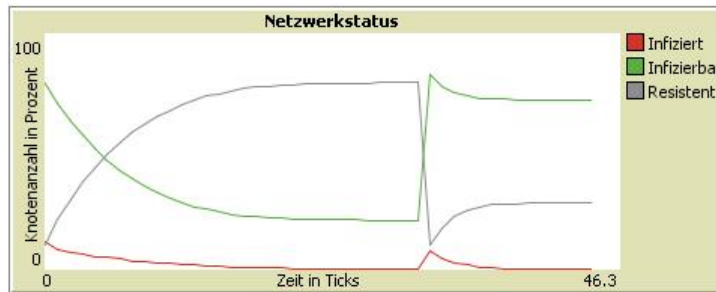


Abbildung 22: Zwei Simulationsdurchläufe mit Thresholdabschaltung (1). In dieser Simulation wurden die gleichen Werte wie bei der reinen SIR-Simulation in Abbildung 21 verwendet. Im ersten Durchgang wurde allerdings zusätzlich ein Threshold von $\Phi = 0,5$ simuliert.

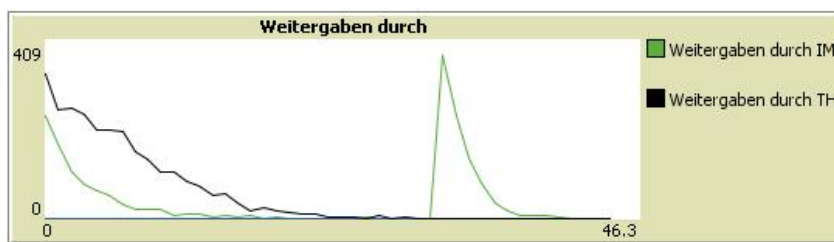


Abbildung 23: Zwei Simulationsdurchläufe mit Thresholdabschaltung (2). Es handelt sich um die gleiche Simulation wie in Abbildung 22. In dieser Grafik sind die Weitergaben durch den Threshold (Weitergaben durch TH) und durch das SIR-Modell (Weitergaben durch IM) einzeln aufgezeigt. Das der Anteil der Weitergaben durch den Threshold größer ist, ist unter anderem dadurch begründet, dass er zuerst berechnet wird.

5 Knotenbeschriftung

Nachdem die Netzwerkerstellung abgeschlossen ist und die Simulationsmethoden bestimmt sind, muss das Netzwerk für die Durchführung vorbereitet werden. Die Knotentypen und anfänglichen Informationsträger müssen bestimmt werden und die für die Simulation erforderlichen Wahrscheinlichkeiten in den Knoten verankert werden.

5.1 Zwei Knotenklassen

Wie bereits in Kapitel 2.4 erläutert, besitzen Knowledge Hubs ein von den restlichen Kommunikationsteilnehmern⁵⁵ verschiedenes Weitergabe- und Annahmeverhalten. Diese Eigenschaft soll auch bei der Simulation Beachtung finden.

Zunächst muss festgelegt werden welcher Knoten als Knowledge Hub und welcher Knoten als Imitator simuliert wird. Vor der Beschriftung⁵⁶ werden alle Knoten als Imitators betrachtet. Nachdem der prozentuale Anteil der Knowledge Hubs vom Benutzer eingegeben wurde, gibt es zwei Möglichkeiten diese zu verteilen. Zum einen können sie rein zufällig über das Netzwerk verteilt werden, zum anderen werden sie vom größten Knotengrad absteigend verteilt. Bei der letzten Variante werden die Knowledge Hubs unter den Social Hubs bestimmt. Diese Knoten entsprechen nun dem klassischen Opinion Leader Begriff, da sie unterschiedliches Weitergabeverhalten und einen großen sozialen Rang in sich vereinen.

Nun können für die zwei Knotentypen unterschiedliche Weitergabe- und Annahmewahrscheinlichkeiten festgelegt werden. Um größtmögliche Flexibilität zu ermöglichen, kann im Tool jede mögliche Weitergabe einzeln bestimmt werden. So können die

⁵⁵ Diese werden weiterhin als Imitators bezeichnet.

⁵⁶ Mit "Beschriftung" wird der Vorgang im Simulationstool beschrieben, den Knotentypen weitere Eigenschaften zu zuordnen. Diese können dann gespeichert und für die Simulation verwendet werden.

Weitergaben von KH⁵⁷ zu IM⁵⁸, von KH zu KH, von IM zu KH und von IM zu IM getrennt bestimmt werden.

Eine solche Trennung in zwei Klassen wurde bereits bei Simulationen mit dem Bass Diffusionsmodell erfolgreich durchgeführt.⁵⁹ Trotz der Unterschiede in den Simulationsmethoden, können die erzielten Werte und Simulationsverläufe eine Hilfe bei der Simulation mit dem entwickelten Simulationstool sein.⁶⁰

5.2 Seeding

Das Seeding ist ein Begriff, der aus dem Marketing stammt und das Platzieren einer Botschaft im Netzwerk beschreibt.⁶¹ Im Simulationstool entspricht dies der Platzierung von Informationsträgern vor dem Start der Simulation. Da die Verteilung dieser Innovators⁶² großen Einfluss auf die Simulation hat, müssen auch hier unterschiedliche Methoden entwickelt werden.

Eine Methode ist das zufällige oder vom Knotengrad abhängige Verteilen, welches bereits bei der Verteilung von Social Hubs Verwendung gefunden hat. Bei dieser Methode werden die Informationsträger weit über das Netzwerk verstreut, was die Verbreitung der Information begünstigt. Dies ist in der Simulation allerdings nicht immer erwünscht, da die Information in der Realität nicht immer beim Seeding über das gesamte Netzwerk verteilt werden kann. Bei einer Verbreitung mittels Massenmedien wäre dies möglich, bei einer Verbreitung an kleinere Zielgruppen hingegen nicht.⁶³

⁵⁷ kurz für Knowledge Hubs

⁵⁸ kurz für Imitators

⁵⁹ vgl. Van den Bulte, Joshi 2006

⁶⁰ Im Bass Diffusionsmodell werden nur "fully connected" Netzwerke simuliert. Das bedeutet, dass jeder Knoten mit jedem Knoten verbunden ist.

⁶¹ vgl. <http://www.viralmarketing.de/2006/10/06/seeding/> Abgerufen am 02.08.2010.

⁶² vgl. Kapitel 2.5

⁶³ Beispielsweise wenn gezielt die Mitglieder eines Fanclubs o.ä. als Innovators gewonnen werden können.

Für den Fall eines gezielten Seedings wird eine weitere Methode implementiert, das Cluster Seeding. Hierbei wird, nachdem der Benutzer die Anzahl der Innovators eingegeben hat, zufällig ein Startknoten ausgewählt. Nachdem dieser Knoten als Innovator markiert wurde, werden seine Nachbarn ausgewählt. Auch diese werden als Innovator markiert, solange die Anzahl der eingegebenen Innovators nicht erreicht ist. Dieser Schritt wiederholt sich solange bis der eingestellte Cluster vollständig ist. Mit dieser Methode werden die Innovators nicht über das Netzwerk verteilt, sondern es entsteht ein zusammenhängender Cluster voller Innovators.

Die Unterschiede in den Seeding Methoden werden deutlich, wenn man die Anzahl der Early Adopter betrachtet.⁶⁴ (Abb. 24)

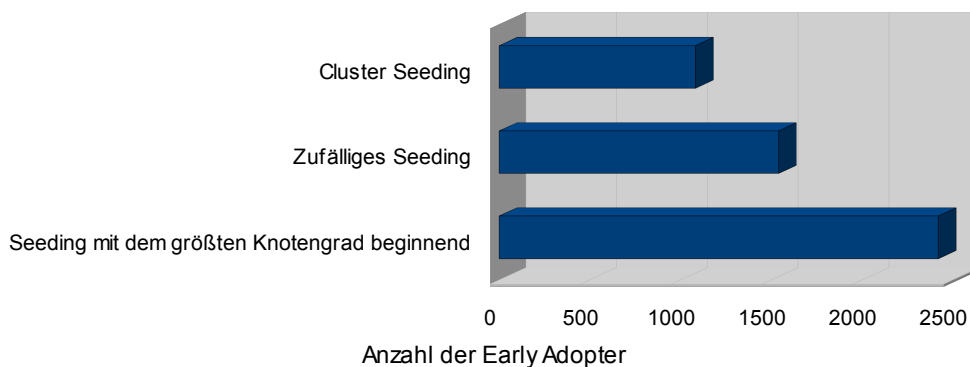


Abbildung 24: Anzahl der Early Adopter nach Seedingmethode. Es wurden 10% Innovators auf 5001 Knoten verteilt. Das Netzwerk ist unbearbeitet. (Ohne CSH oder Megahubs). In einem bearbeiteten Netzwerk ist noch mit deutlicheren Unterschieden zu rechnen.

5.3 Weitergabe der Innovators

Eine Methode, die im Marketing immer mehr Verbreitung findet, ist das Rekrutieren von "Agents", die eine Verbreitung beschleunigen sollen. Diese werden als Innovators ausgewählt und verbreiten die Information deutlich stärker als normale

⁶⁴ vgl. Kapitel 2.5

Kommunikationsteilnehmer.⁶⁵ Die Gründe dafür sind vielfältig. Die Wirkung ist hingegen einfach und kann in der Simulationslösung umgesetzt werden. Dazu kann die Weitergabewahrscheinlichkeit von Innovators separat bestimmt werden. Ein separates Annahmeverhalten ist hingegen nicht nötig, da sich verschiedenste Kommunikationstypen unter den "Agents" befinden.

⁶⁵ vgl. Walker (2004)

6 Beispielsimulation

Nachdem alle Faktoren und Parameter für die Simulation bestimmt sind, ist es möglich, diese anhand eines Beispiels näher zu erläutern. Dabei wird ein hypothetischer Fall simuliert und ausgewertet.

6.1 Ausgangslage

Gegenstand des Beispiels sei das hypothetische Unternehmen X, das für die Einführung eines neuen Produktes einen Newsletter verbreiten will. Im Newsletter wird das Produkt mit einem humoristischen Bild vorgestellt und die Empfänger werden aufgefordert den Newsletter an Bekannte und Freunde weiterzuleiten. Diese Ausgangslage entspricht den Simulationsmöglichkeiten des entwickelten Simulationstools. Allerdings müssen zunächst die benötigten Parameter ermittelt werden.

6.1.1 Netzwerkerstellung

Die Firma X hat durch eine Analyse ihrer Zielgruppe bestimmt, dass die Struktur gegenüber anderen Kommunikationsnetzen im Internet keine Besonderheiten aufweist. Es gibt weder Mega Hubs noch sind die Kunden besonders gleichmäßig vernetzt. Daher wird für die Simulation ein Abweichparameter (a_{Abweich})⁶⁶ von 0 angenommen. Nach der Erzeugung eines Netzwerkes mit 20001 Knoten wird zusätzlich noch die Funktion CSH mit den Werten

$[(+0)(+8)(+8)(+2)]$ ⁶⁷ durchgeführt um die Anzahl der Social Hubs anzugleichen. Nach einem Vergleich mit der Zielgruppe zeigt sich, dass bei einem entstanden größten Knotengrad von 56 keine Obergrenze für Social Hubs notwendig ist.⁶⁸ Das erstellte Netzwerk stellt keine exakte Kopie des von der Firma X untersuchten

⁶⁶ vgl. Kapitel 3.4

⁶⁷ Werte aufsteigend nach Gruppen angegeben; vgl. Kapitel 3.5.3

⁶⁸ vgl. Kapitel 3.6

Netzwerkes dar, aber eine für den Rahmen einer Simulation ausreichende Näherung. (Abb. 25)

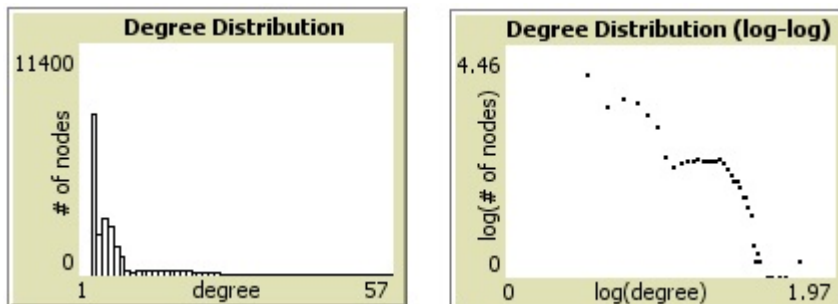


Abbildung 25: Knotengradverteilung des Beispielnetzwerkes. Sie entspricht annähernd den Werten des Viral Manager aus Kapitel 3.5.3.

Weiter hat die Firma X bestimmt, dass näherungsweise die größten 13% der Netzwerkelemente als Knowledge Hubs bestimmbar sind. Diese 13% stellen eine Verbindung von Social Hub und Knowledge Hub dar und entsprechen dem klassischen Opinion Leader Begriff. Die größten Social Hubs, die nicht gleichzeitig auch Knowledge Hubs sind, haben nun den Knotengrad von 9.

6.1.2 Weitergabeparameter

Nachdem die Firma X ausführliche Tests und Marktforschung mit dem Newsletter betrieben hat, können folgende Weitergabe- und Annahmeparameter bestimmt werden: (Tab. 1)⁶⁹

⁶⁹ Dabei ist klar, dass die Bestimmung aller Parameter und die Unterteilung in Kommunikationstypen in der Realität nicht so eindeutig möglich ist. In diesem hypothetischen Beispiel allerdings ist jeder Wert genau bestimmt. Gleiches gilt für die Netzwerkerstellung.

Parameter	Wert
Weitergabewahrscheinlichkeit KH (Knowledge Hub)	50%
Weitergabewahrscheinlichkeit IM (Imitator)	50%
Weitergabewahrscheinlichkeit IN (Innovator)	50%
Annahmewahrscheinlichkeit KH zu IM	10%
Annahmewahrscheinlichkeit KH zu KH	35%
Annahmewahrscheinlichkeit IM zu IM	20%
Annahmewahrscheinlichkeit IM zu KH	5%
Threshold	67%

Tabelle 1: Ausgangswerte der Beispielsimulation.

6.2 Beispielseeding

Zunächst werden die unterschiedlichen Methoden des Seeding simuliert. Dabei wird untersucht wie mit dem kleinsten Aufwand, d.h. nur ein informationstragender Knoten zu Beginn der Simulation, die beste Ausbreitung erzielt werden kann. (Abb. 26 und Tab. 2)

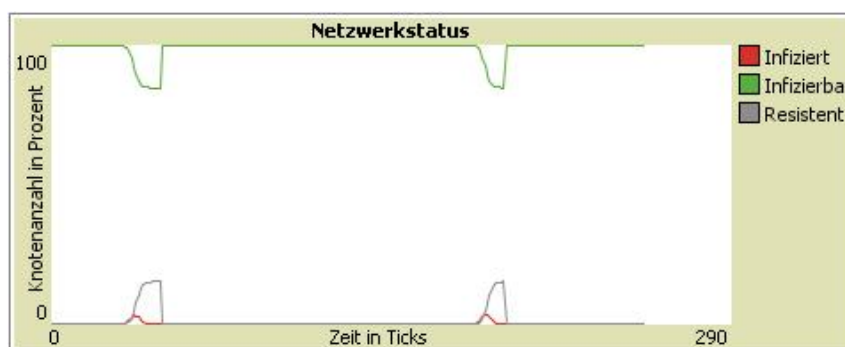


Abbildung 26: Informationsausbreitungen. Untersucht werden Informationsausbreitungen, die durch lange Kettenreaktionen (Kaskaden) gekennzeichnet sind. Sie unterscheiden sich in der Regel deutlich in Größe und Dauer von schnell versiegenden Ausbreitungen. In diesem Beispiel traten unter 100 Versuchen nur zwei relevante Informationsausbreitungen auf.

Beispielsimulation

Versuchsnummer	Knoten-typ	Knoten-grad	Ausbreitungswahrscheinlichkeit
1	Imitator	3	2,8%
2	Imitator	9	7,9%
3	Knowledge Hub	15	32,9%
4	Knowledge Hub	56	48,9%

Tabelle 2: Beispielseeding mit Einzelknoten.⁷⁰

Der Knotengrad des Innovators hat einen deutlichen Einfluss auf die Ausbreitungswahrscheinlichkeit der Information. Bei dem Versuch treten allerdings auch Besonderheiten auf. Zum einen ist der Sprung von Versuch 2 zu 3 sehr groß. Obwohl sich der Knotengrad nicht einmal verdoppelt, ist die Auswirkung auf die Ausbreitungswahrscheinlichkeit sehr stark. Zum anderen hatte die Informationsausbreitung bei allen vier Seedingmethoden stets die Größe von etwa 15,4% des Gesamtnetzwerkes. Sie ist also unabhängig vom Innovator. Die Erklärung für diesen Effekt findet sich im typischen Ausbreitungsverlauf. (Abb. 26)

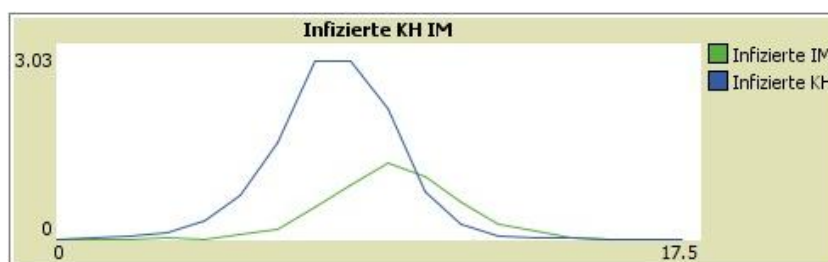


Abbildung 27: Typischer Ausbreitungsverlauf der Beispielsimulation.

Am Verlauf ist zu erkennen, dass sich die Information größtenteils unter den Knowledge Hubs ausbreitet. Insgesamt werden fast die gesamten 13% Knowledge Hubs infiziert, was gleichzeitig

⁷⁰ Alle Daten der Beispielsimulation werden als Durchschnittswerte von 1000 Simulationsdurchläufen erhoben.

bedeutet, dass nicht einmal 3% der Imitators infiziert werden. Dieser typische Verlauf ändert sich auch bei unterschiedlichen Innovators nicht und lässt den Schluss zu, dass er durch andere Netzwerkparameter bestimmt wird. Damit lässt sich neben der relativ konstanten Ausbreitungsgröße auch der Sprung in der Ausbreitungswahrscheinlichkeit bei Versuch 2 zu Versuch 3 erklären. Wenn sich die Information fast ausschließlich unter den Knowledge Hubs ausbreitet, ist es für die Verbreitung günstig, wenn beim Start bereits einer von ihnen Informationsträger ist. Wenn nicht, muss die Information erst noch einen der Knowledge Hubs erreichen und infizieren.

Nun soll die Auswirkung der Weitergabewahrscheinlichkeit der Innovators untersucht werden. Dazu wird sie in den Versuchen 1 bis 4 schrittweise auf 100% angehoben und die Wahrscheinlichkeit der Informationsausbreitung ermittelt. (Abb. 28)

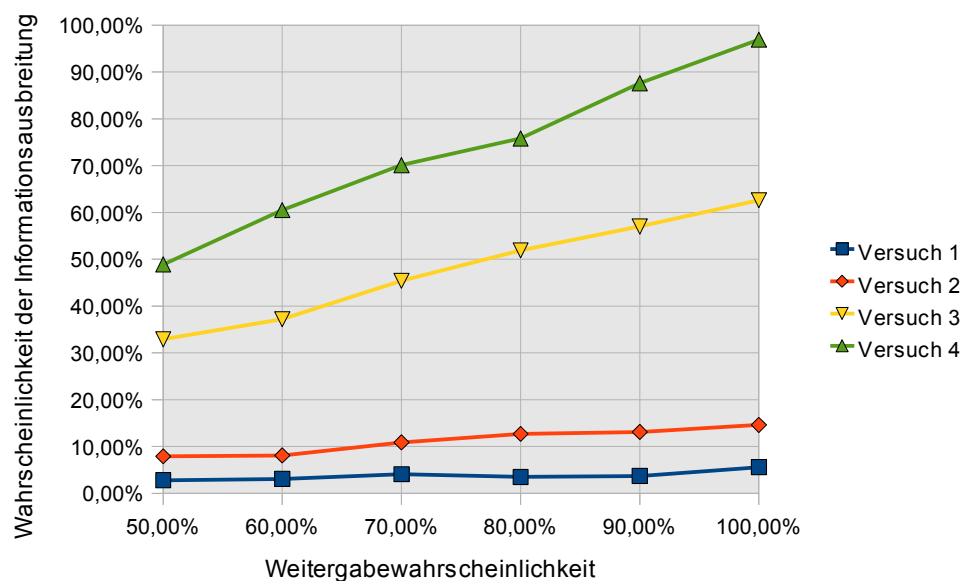


Abbildung 28: Auswirkung der Weitergabewahrscheinlichkeit der Innovators.

Die Wahrscheinlichkeit der Informationsausbreitung steigt bei zunehmender Weitergabezahl an. Dies allein ist noch keine

überraschende Erkenntnis, aber auch hier lassen sich zwei interessante Fakten aufzeigen. Zum einen gelingt es selbst dem größten Knoten, wenn er zu 100% seine 56 Kontakte nutzt, nicht in jedem Fall eine große Informationsausbreitung auszulösen. Zum anderen hat der Knoten aus dem Versuch 3 bei hohen Weitergabewahrscheinlichkeiten eine größere Chance Informationsausbreitungen auszulösen, als der deutlich vernetztere Knoten aus Versuch 4 mit geringen Weitergabewahrscheinlichkeiten. (Tab. 3)

Knoten Versuch 3		Knoten Versuch 4	
Absolute Weitergaben (Weitergabewahrscheinlichkeit)	Ausbreitungswahrscheinlichkeit	Absolute Weitergaben (Weitergabewahrscheinlichkeit)	Ausbreitungswahrscheinlichkeit
15 (100%)	62,6%	28 (50%)	48,9%
		33,6 (60%)	60,5%

Tabelle 3: Weitergaben und Ausbreitungswahrscheinlichkeit im Vergleich.

Eine Erklärung für diese Auffälligkeit zu finden, gestaltet sich deutlich schwieriger und würde über den Rahmen dieser Arbeit hinaus gehen. Es wäre möglich, dass es wichtig ist, dass die Knoten alle ihre Kontakte infizieren um eine zusammenhängende Menge infizierter Knoten zu bilden. Oder dass der größte Knoten nicht so wichtig ist, da er aufgrund seiner vielen Kontakte höchstwahrscheinlich sowieso mit der Information in Kontakt gerät.

6.3 Weitergabewahrscheinlichkeit

Im nächsten Versuch werden die Auswirkungen der Weitergabewahrscheinlichkeiten von Knowledge Hubs und Imitators simuliert. Dabei wird die Wahrscheinlichkeit schrittweise von 50% auf 70% angehoben ohne die Weitergabewahrscheinlichkeit der Innovators

zu verändern. Im Seeding werden 10 zufällige Imitators mit dem Knotengrad 9 als Informationsträger bestimmt. (Abb. 29)

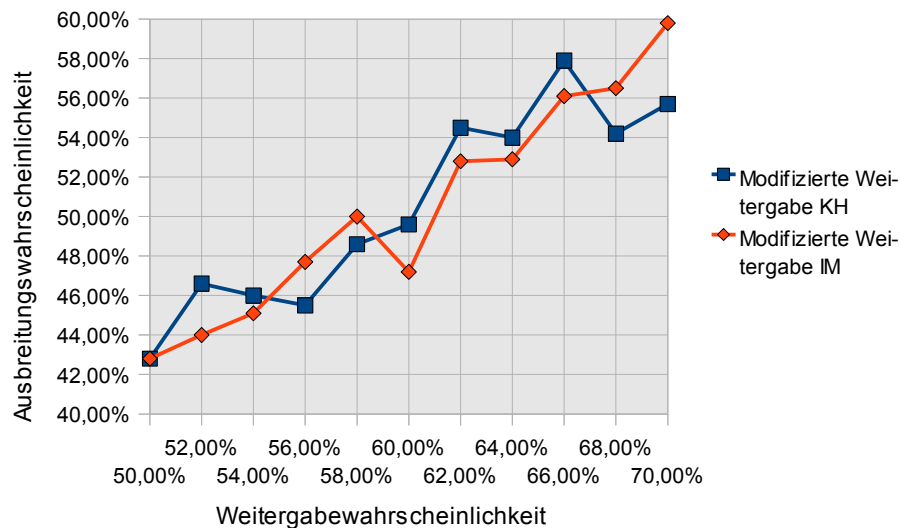


Abbildung 29: Auswirkungen der Weitergabewahrscheinlichkeit auf die Informationsausbreitung.

Es ist eine deutliche Steigerung der Anzahl der Informationsausbreitungen zu verzeichnen. Im simulierten Rahmen kann allerdings nicht bewertet werden, welcher Parameter bei dieser Variante die größere Auswirkung auf die Ausbreitungswahrscheinlichkeit hat. Die Auswirkung auf die Ausbreitungsgröße und Ausbreitungsdauer ist besser ersichtlich. (Abb. 30 und Abb. 31)

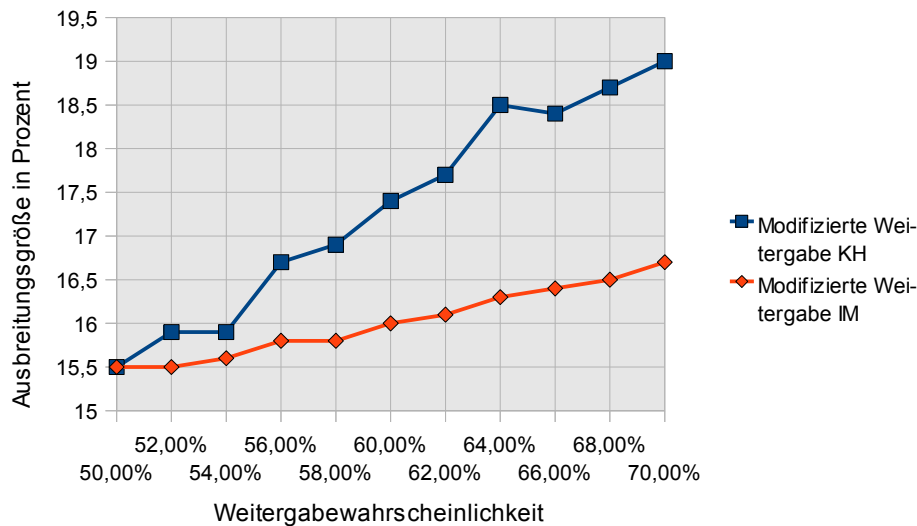


Abbildung 30: Auswirkungen der Weitergabewahrscheinlichkeit auf die Ausbreitungsgröße.

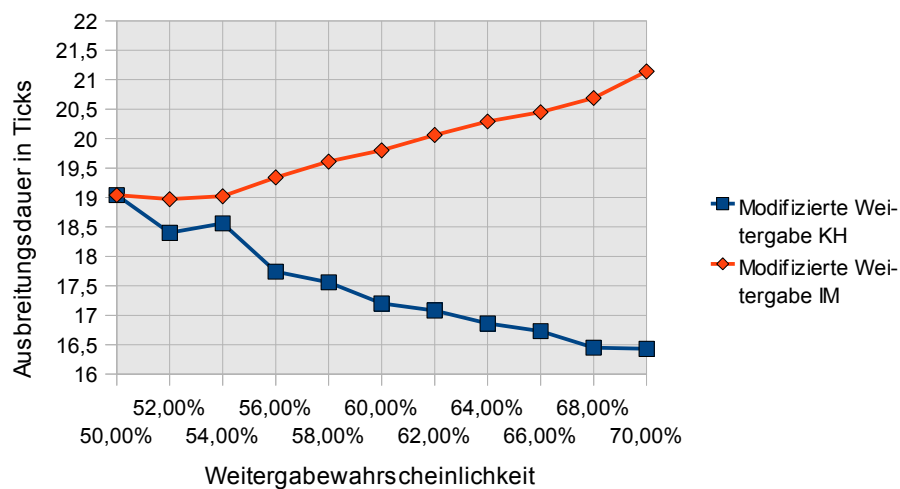


Abbildung 31: Auswirkungen der Weitergabewahrscheinlichkeit auf die Ausbreitungsdauer.⁷¹

An den Auswertungen lässt sich ablesen, dass die erhöhte Weitergabe von Knowledge Hubs einen deutlich größeren Einfluss auf

⁷¹ Ein Tick ist eine Zeiteinheit in der Simulationsumsetzung in Netlogo. Nachdem alle informationstragenden Knoten berechnet wurden, wird ein Tick gezählt und die Berechnung erfolgt von neuem.

die Ausbreitungsgröße hat, als die Erhöhung der Weitergabe der Imitators. Bedenkt man die Netzstruktur und die Verteilung der Knowledge Hubs unter den Social Hubs, erklärt sich diese Beobachtung. Denn die Knowledge Hubs haben deutlich mehr Kontakte als die Imitators. Daher bedeutet die gleiche prozentuale Änderung der Weitergabewahrscheinlichkeiten von Knowledge Hubs und Imitators einen größeren Anstieg an absoluten Weitergaben bei den Knowledge Hubs.

Die Ausbreitungsdauer fällt bei einer Anhebung der Weitergabewahrscheinlichkeit der Knowledge Hubs deutlich. Mit der Beachtung des typischen Ausbreitungsverlaufes ist auch diese Beobachtung zu erklären. Denn die Information breitet sich zum Großteil unter den Knowledge Hubs aus. Daher erreicht die Ausbreitung auch schneller die Grenze der möglichen Weitergabepartner unter den Knowledge Hubs. Bei der Anhebung der Weitergabewahrscheinlichkeit der Imitators steigt die Ausbreitungsdauer. Diese Beobachtung lässt sich damit erklären, dass sich die Information etwas besser unter den Imitators verbreitet und diese Verbreitung mangels vieler Kontakte langsamer verläuft.

6.4 Annahmewahrscheinlichkeit

Ein weiterer zentraler Bestandteil der Simulation sind die Annahmewahrscheinlichkeiten. Daher sollte auch ihre Auswirkungen einzeln betrachtet werden.

6.4.1 Knowledge Hubs zu Imitators

Im ersten Versuch wird die Annahmewahrscheinlichkeit verändert, die den Imitators zugeordnet wurde, wenn sie die Information von Knowledge Hubs erhalten. Sie wird schrittweise um 20 Prozentpunkte auf 30% angehoben, wobei das Seeding immer noch mit 10 zufälligen Imitators mit dem Knotengrad 9 durchgeführt wird. (Abb. 32)

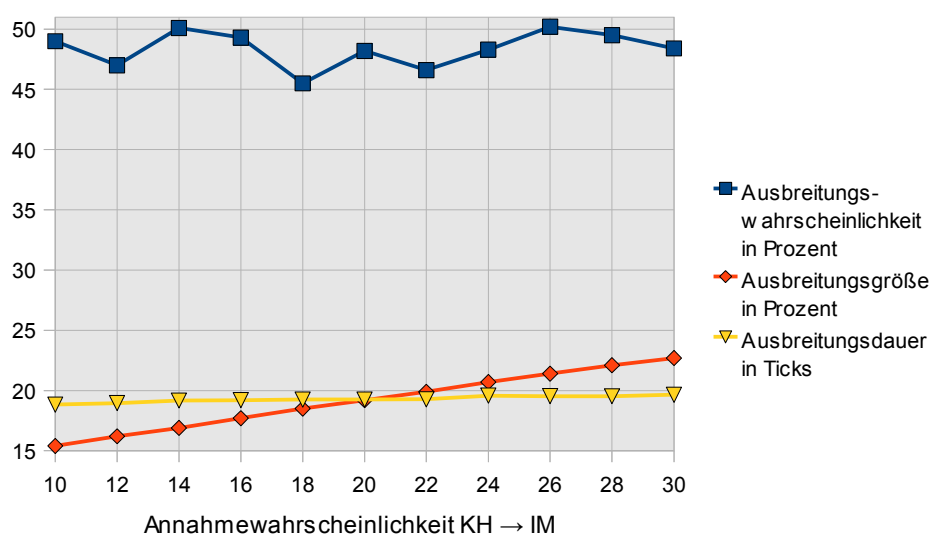


Abbildung 32: Auswirkungen der Annahmewahrscheinlichkeit KH → IM.

Deutlich ist zu erkennen, dass die Ausbreitungswahrscheinlichkeit nicht signifikant ansteigt. Dies ist aufgrund der typischen Verbreitung unter den Knowledge Hubs auch nicht anders zu erwarten. Erstaunlich ist, dass obwohl die Ausbreitungsgröße sich um fast 50% erhöht, die Ausbreitungsdauer nur minimal steigt. Dies legt die Vermutung nahe, dass die bessere Infizierung der Imitators durch Knowledge Hubs aufgrund ihrer vielen Kontakte zwar effektiv ist, aber die Information in diesem Beispiel nicht lange unter den Imitators "überlebt".

6.4.2 Knowledge Hubs zu Knowledge Hubs

Wie im vorangegangenen Versuch wird das Seeding beibehalten und die Annahmewahrscheinlichkeit um 20 Prozentpunkte angehoben. (Abb. 33)

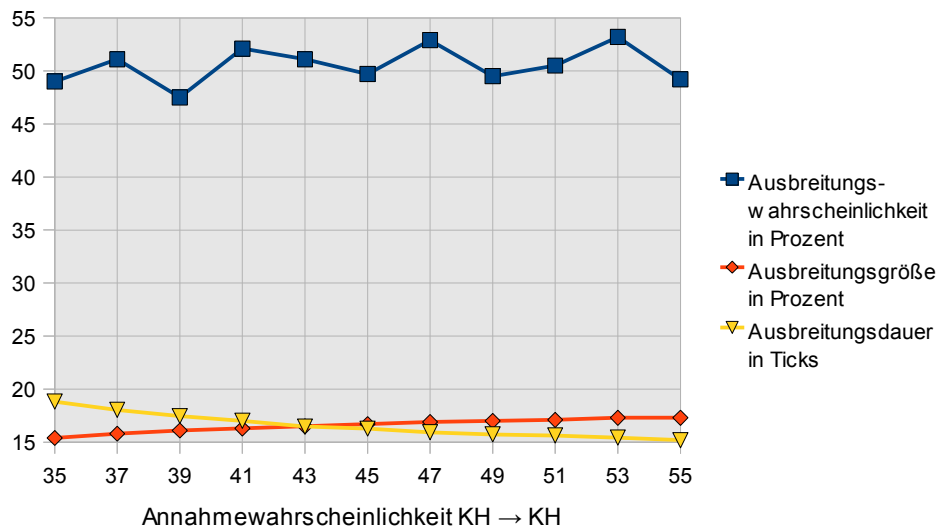


Abbildung 33: Auswirkungen der Annahmewahrscheinlichkeit KH → KH.

Obwohl in diesem Versuch die Ausbreitungswahrscheinlichkeit etwas höher liegt als im vorangegangenen Versuch, ist auch hier kein deutlicher Anstieg zu beobachten. Anscheinend ist die notwendige "Schwelle", die für eine Ausbreitung unter den Knowledge Hubs nötig ist, bereits überschritten und weitere Anhebung der Annahmewahrscheinlichkeit unter den Knowledge Hubs bringt nur noch kleine Erfolge. So wird Ausbreitungsdauer gesenkt und die Ausbreitungsgröße leicht angehoben.

6.4.3 Imitators zu Imitators

Der gleiche Versuch wird erneut durchgeführt. Allerdings wird nun die Annahmewahrscheinlichkeit betrachtet, wenn Imitators die Informationen von anderen Imitators erhalten. (Abb. 34)

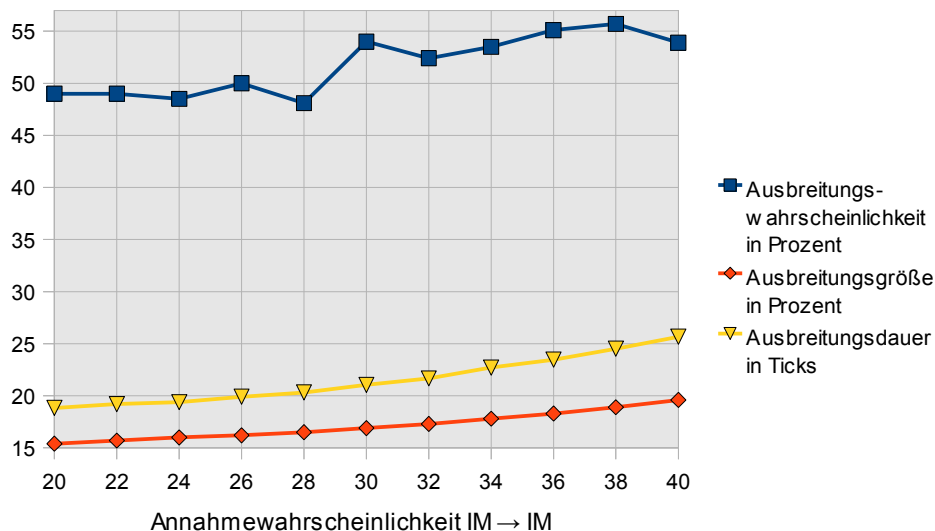


Abbildung 34: Auswirkungen der Annahmewahrscheinlichkeit IM → IM.

Zum ersten Mal ist bei der Modifizierung der Annahmeparameter eine klare Steigerung der Ausbreitungswahrscheinlichkeit zu beobachten. Mit einem Blick auf die Ausbreitungsdauer lässt sich auch dieses Phänomen erklären. Bleibt die Information länger im Netzwerk, indem sie unter der breiten Masse der Imitators weitergegeben wird, steigt auch die Wahrscheinlichkeit, dass sie irgendwann eine ausreichende Menge an Knowledge Hubs erreicht. Die Ausbreitungsgröße steigt dabei nur moderat an, da die Imitators nicht über so viele Kontakte wie die Knowledge Hubs verfügen.

6.4.4 Imitators zu Knowledge Hubs

Nun wird der gleiche Versuch auch mit dem letzten Annahmeparameter durchgeführt. (Abb. 35)

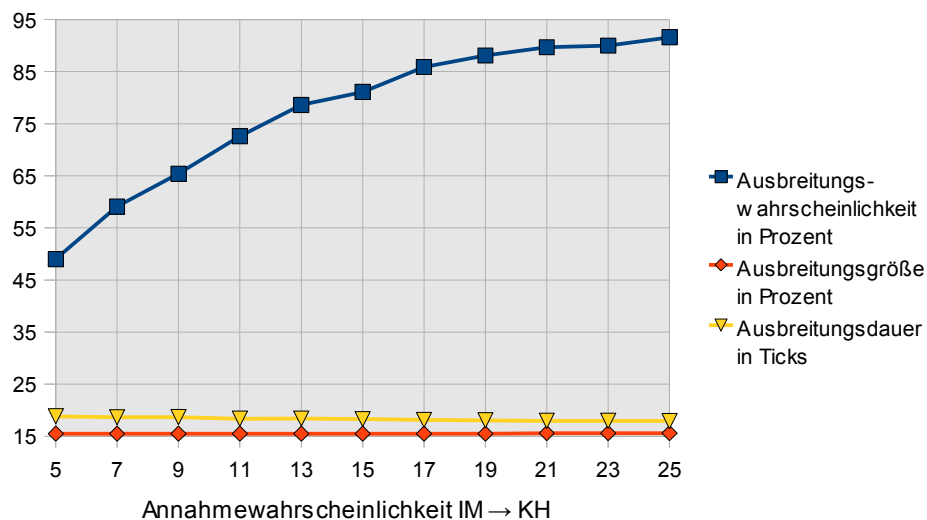


Abbildung 35: Auswirkungen der Annahmewahrscheinlichkeit IM → KH.

In diesem Versuch ist die mit Abstand deutlichste Steigerung der Ausbreitungswahrscheinlichkeit zu beobachten. Wenn man bedenkt, dass beim Seeding nur Imitators infiziert werden und diese dann erst die Information zu den Knowledge Hubs tragen müssen, ist die starke Steigerung nicht verwunderlich. Interessant ist ebenfalls, dass die Ausbreitungsgröße praktisch unverändert bleibt und die Ausbreitungsdauer nur minimal fällt.

6.5 Netzwerkstruktur

Im letzten Versuch der Beispielsimulation soll die Auswirkung der Netzwerkstruktur auf die Simulationsergebnisse angedeutet werden. Dabei wird das bereits ausführlich beleuchtete Netzwerk mit zwei neuen Varianten verglichen, bei denen nur die Netzwerkstruktur verändert wird, aber die Weitergabeparameter beibehalten werden.

Beispielsimulation

Das erste neu erstellte Versuchsnetzwerk 2 entspricht dem bisher verwendeten Netzwerk, bis auf die Tatsache, dass kein CSH durchgeführt wurde. (Abb. 36)

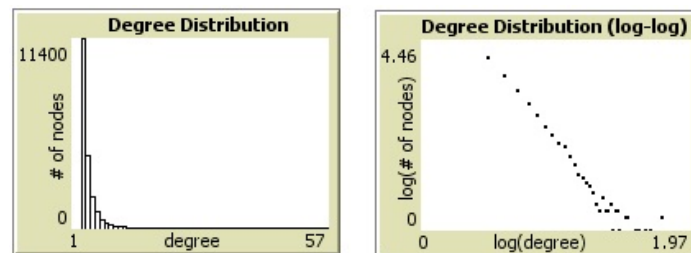


Abbildung 36: Knotengradverteilung des Versuchsnetzwerkes 2.

Das Seeding wird wie in den letzten Versuchen durchgeführt, d.h. dass zufällig aus den größten Imitators 10 Innovators ausgewählt werden. Da aber ohne CSH deutlich weniger Knoten mit großen Knotengraden existieren, haben die Innovators nur noch einen Knotengrad von 5 anstatt von 9. Die Ergebnisse der Simulation unterscheiden sich deutlich von den Ergebnissen des bisherigen Netzwerkes. (Abb. 37)

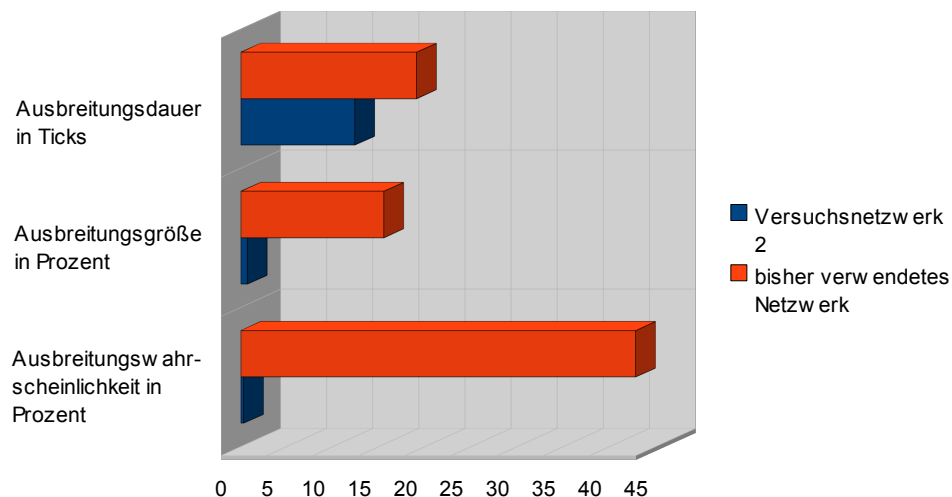


Abbildung 37: Vergleich mit Versuchsnetzwerk 2.

Es kommt im Versuchsnetzwerk 2 nun zu wesentlich weniger Informationsausbreitungen (0,3%) mit einer deutlich geringeren Größe (0,67%). Neben dem schlechteren Seeding ist die Ursache darin zu finden, dass nun die Knowledge Hubs bedeutend schlechter vernetzt sind. Die Information findet keine Wege mehr alle von ihnen zu erreichen. Erstaunlicherweise verweilt die Information trotz der geringen Ausbreitungsgröße relativ lang im Netzwerk. Mit diesem Extremfall wird sehr gut verdeutlicht, dass nicht immer ein Zusammenhang zwischen Ausbreitungsdauer und Ausbreitungsgröße besteht. Die Information kann relativ lang im Netzwerk zirkulieren ohne jemals eine relevante Ausbreitung auszulösen.

Die zweite Variante, die simuliert werden soll, ist die des Mega Hubs. Dabei wird erneut ein Netzwerk erstellt, bei dem mittels eines Abweichparameters und einer Obergrenze ein Mega Hub mit dem Knotengrad 1000 erzeugt wurde. Danach wird wieder die Funktion CSH mit den selben Parametern wie beim Ausgangsnetzwerk durchgeführt. (Abb. 38)

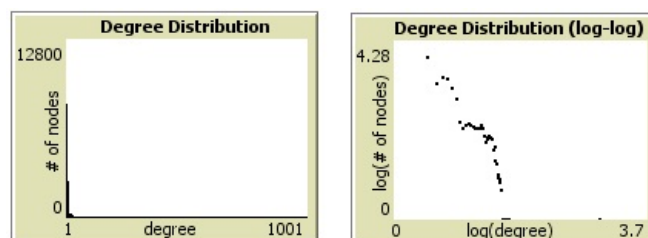


Abbildung 38: Knotengradverteilung des Versuchsnetzwerkes 3.

Bei dieser Netzstruktur bietet sich an von der bisher verwendeten Seedingmethode abzuweichen. So wird nun der Mega Hub als Innovator genutzt. (Abb. 39)

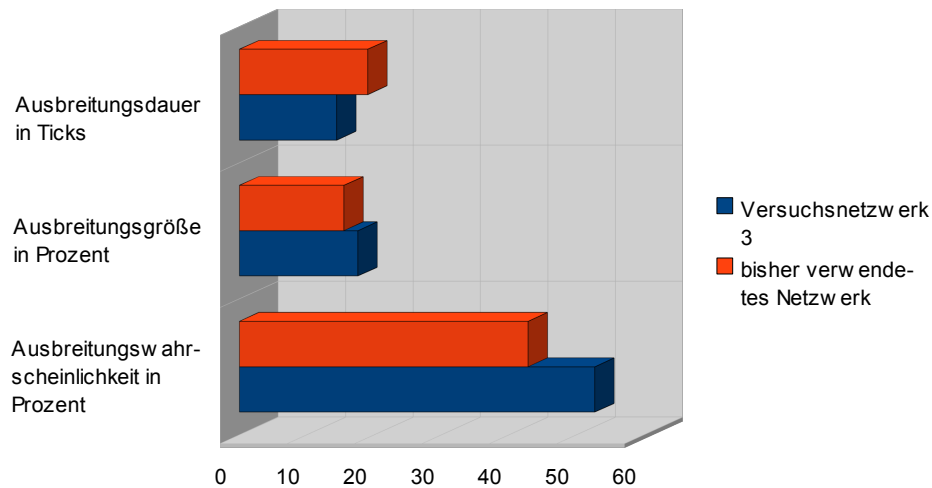


Abbildung 39: Vergleich mit Versuchsnetzwerk 3.

In der Simulation kommt es nicht zu extrem großen Abweichungen bezüglich Ausbreitungswahrscheinlichkeit, -dauer oder -größe. Dies hätte man aufgrund des großen Knotengrades durchaus erwarten können. Eine mögliche Erklärung wäre, dass es auch dem Mega Hub nicht gelingt die typische Verteilung zu durchbrechen.⁷²

6.6 Fazit der Beispielsimulation

Aus den vorangegangenen Simulationen lassen sich, neben den Möglichkeiten die das Simulationstool bietet, auch erste Einblicke in die Informationsausbreitung gewinnen. Zum Beispiel wie stark der Ausbreitungsverlauf von den Netzwerkparametern abhängt. Selbst bei tausendfacher Ausführung tritt fast immer ein relativ ähnlicher Verlauf auf. Ebenfalls lässt sich erkennen wie Aufgrund des sehr dynamischen Ausbreitungsprozesses die Ausbreitungswahrscheinlichkeit schwankt. Um weitere Erkenntnisse über die Verbreitung von Informationen in Netzwerken gewinnen zu können, müssen weitere Simulationen durchgeführt werden. Interessant wäre unter anderem eine Veränderung des Threshold, oder

⁷² vgl. Abbildung 27

eine Simulation mit unterschiedlichen Innovators gleichen Knoten-grades. Auch durch eine genauere Betrachtung des Zeitraumes kurz bevor sich eine Information mit einer Kettenreaktion verbreitet, könnte zur Gewinnung neuer Einsichten beitragen.

Abschließend stellt sich die Frage, welche Schlüsse die Firma X aus den bisher durchgeführten Simulationen ziehen kann. Die wohl sicherste Erkenntnis ist, dass sich der Newsletter wahrscheinlich nur unter den Knowledge Hubs verbreiten wird. Andere Erkenntnisse wie Ausbreitungswahrscheinlichkeiten oder Reichweiten sind stark von der Genauigkeit der ermittelten Parameter und des Netzwerkes abhängig und somit nur als Richtlinie interpretierbar.

7 Fazit

Wie in der Beispielsimulation im Kapitel 6 bereits angedeutet wurde, kann mit dem Simulationstool eine Vielzahl an Ausbreitungseigenschaften ermittelt werden. Ob sich aus diesen Eigenschaften anwendbare Verbesserungsmöglichkeiten für die Ausbreitung entwickeln lassen, konnte noch nicht überprüft werden, lag aber auch nicht im Fokus der Arbeit. Das Ziel, ein Tool zu erstellen, das auch ohne tiefgründiges Hintergrundwissen bedienbar ist und eine flexible Berücksichtigung zielgruppenspezifischer Werte ermöglicht, wurde dennoch erreicht.

Unterstützt wird die Simulation durch eine in Echtzeit mitlaufende Ausgabe der Netzwerkparameter und eine einfache Bearbeitung der Simulationsergebnisse mit Tabellenkalkulationsprogrammen. Damit sind die Grundlagen für weitere Simulationen geschaffen worden, auf die in Zukunft aufgebaut werden kann.

Ein wichtiger Punkt ist die Richtigkeit der Ergebnisse des Tools. Es stellt sich die Frage, ob die Ergebnisse trotz der Annahmen und Verallgemeinerungen, die auf dem Weg zum fertigen Tool getroffen wurden, korrekt sind.⁷³ Wenn auch die Richtigkeit des Simulationstools schwer zu bewerten ist, kann festgestellt werden, dass sich die Ergebnisse aller Versuche mit den typischen Ausbreitungskurven bisheriger Simulationsansätze decken. So folgt der Graph mit der Anzahl der resistenten Knoten einem für Epidemien typischen S-Kurvenverlauf.⁷⁴ Auch die Auswertung nach Knowledge Hubs und Imitators gleicht den Ergebnissen ähnlich konzipierter Simulationen.⁷⁵ Diese Vergleiche zeigen, dass die Ergebnisse, die mit dem Simulationstool erzielt werden können durchaus wissenschaftlichen Vergleichen standhalten.

Dabei bietet die Arbeit mit Netlogo und den ABM neue Möglichkeiten der Analyse der Informationsausbreitung. Es ist nun prinzipiell möglich jedes Element in der Mikro Ebene zu beobachten und neue Schlüsse aus emergenten Effekten in der Makro

⁷³ zumindest so korrekt, dass Aussagen über die Informationsausbreitung stichhaltig untermauert werden können.

⁷⁴ vgl. Watts 2003, 172

⁷⁵ vgl. Abb. 27, vgl. Van den Bulte, Joshi 2006

Fazit

Ebene zu ziehen. Des Weiteren ist es durch die niedrige Lernkurve von Netlogo für Interessierte relativ leicht möglich Modifizierungen wie ein anderes Entscheidungssystem, eine andere Netzwerkstruktur oder andere Auswertungsprozeduren zu entwickeln.

Literaturverzeichnis

Fachbücher

Barabási, Albert-László: Linked. How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life. New York 2003

Gladwell, Malcolm: Tipping Point. Wie kleine Dinge Großes bewegen können, 5. Auflage. München 2002

Granovetter, Mark: The strength of weak ties: a network theory revisited. In: Sociological Theory Volume 1 (1983), Seiten 201-233

Klinger, Michaela: Virales Marketing. Die Macht der sozialen Netzwerke. Saarbrücken 2006

Rogers, Everett M.: Diffusion of Innovations. 4. Auflage. New York 1995

Rosen, Emanuel: The Anatomy of Buzz Revisited. Real-life Lessons in Word-of-Mouth Marketing. 1. Auflage. New York 2009

Rürup, Matthias: Innovationswege im deutschen Bildungssystem. Die Verbreitung der Idee "Schulautonomie" im Ländervergleich. 1. Auflage. Wiesbaden 2007

Watts, Duncan J.: Six Degrees. The Science of a Connected Age. 1. Auflage. 2003

Hochschulschriften

Herman, Frederik: Virales Marketing. Karlsruhe 2004 <http://www.slideshare.net/netzkobold/virales-marketing-diplomarbeit-von-frederik-hermann-presentation>. Abgerufen am 22.07.2010

Sonstige Schriften

Horvitz, Eric; Leskovec Jure: Planetary-Scale Views on a Large Instant-Messaging Network. 2008 http://research.microsoft.com/en-us/um/people/horvitz/leskovec_horvitz_www2008.pdf Abgerufen am 24.07.2010

Percy, Katrina(Hrsg): Viral Manager. How users Share Viral Content Online. Ohne Veröffentlichungsdatum. http://www.viralmanager.co.uk/strategy/research_documents/How_Users_Share_Viral_Content_Online. Abgerufen am 22.07.2010

Van den Bulte, Christophe; Joshi, Yogesh V.: New Product Diffusion with Influentials and Imitators. <http://knowledge.wharton.upenn.edu/papers/1322.pdf> Abgerufen am 02.08.2010

Walker, Rob: The Hidden (in Plain Sight) Persuaders 05.12.2004. <http://www.nytimes.com/2004/12/05/magazine/05BUZZ.html>. Abgerufen am 29.07.2010

Wilensky, U. (2005). NetLogo Preferential Attachment model. <http://ccl.northwestern.edu/netlogo/models/PreferentialAttachment>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (Quellenangabe nach den Copyright Vorschriften des Verfassers)

Wilensky,U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (Quellenangabe nach den Copyright Vorschriften des Verfassers)

Page, Scott E. "agent-based models." The New Palgrave Dictionary of Economics. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, 2008. The New Palgrave Dictionary of Economics Online. Palgrave Macmillan. 06 August 2010 <http://www.dictionaryofeconomics.com/article?id=pde2008_A000218> doi:10.1057/9780230226203.0016 (Quellenangabe nach den Copyright Vorschriften des Verfassers)

Zeitschriften

Albert, Réka; Barabási, Albert László: Statistical mechanics of complex networks. In: Review of modern Physics. Volume 74 (2002), Seiten 47-97

Chan, Kenny; Shekhar, Misra: Characteristics of the Opinion Leader: A new dimension. In: Journal of Advertising, Volume 10 (2000) Issue 3, Seiten 53-60.

Kermack W.O.; McKendrick A.G (Hrsg): Contributions to the mathematical theory of epidemics (Part I), In: Proceedings of the Royal Society, Nr.115A (1927)

Latané, Bibb; L'Herrou Todd: Spatial Clustering in the Conformity Game: Dynamic Social Impact in Electronic Groups. In: Journal of Personality and Social Psychology. Volume 70 (1996), Issue 6. Seiten 1218-1230

Watts, Duncan J.: A simple model of global cascades on random networks. In: PNAS, Volume 99 (2002) no.9

Watts, Duncan J.; Dodds, Peter Sheridan: Influentials, Networks, and Public Opinion Formation. In: Journal of Consumer Research, Volume 34 (2007), Issue 4, Seiten 441-458

Anhang

Vorwort zu den Kurzanleitungen

Das Simulationstool ist für eine bessere Übersicht nach den Erstellungsschritten dreigeteilt. So wird die Netzwerkerstellung, die Netzwerkbeschriftung und die Simulation in einem einzelnen Tool durchgeführt. Daher muss das Netzwerk nach der Bearbeitung mit einem der Tools stets gespeichert und mit dem nächsten geladen werden.

Das Tool wurde mit der Version 4.1 mit Netlogo erstellt und getestet. Da die Tools nicht allein lauffähig sind, muss zunächst das Programm Netlogo von der offiziellen Homepage heruntergeladen und installiert werden.⁷⁶ Danach können die Tools mit einem Doppelklick gestartet werden.

Es ist allgemein zu empfehlen das Darstellungsfenster durch das Entfernen des Hakens bei "View Updates" zu deaktivieren. Dadurch ist ein deutlicher Performancegewinn spürbar, zumal bei großen Netzwerken keine sinnvolle Darstellung möglich ist.

⁷⁶ <http://ccl.northwestern.edu/netlogo/download.shtml> abgerufen am 24.08.2010

Kurzanleitung für das Tool zur Netzwerkerstellung

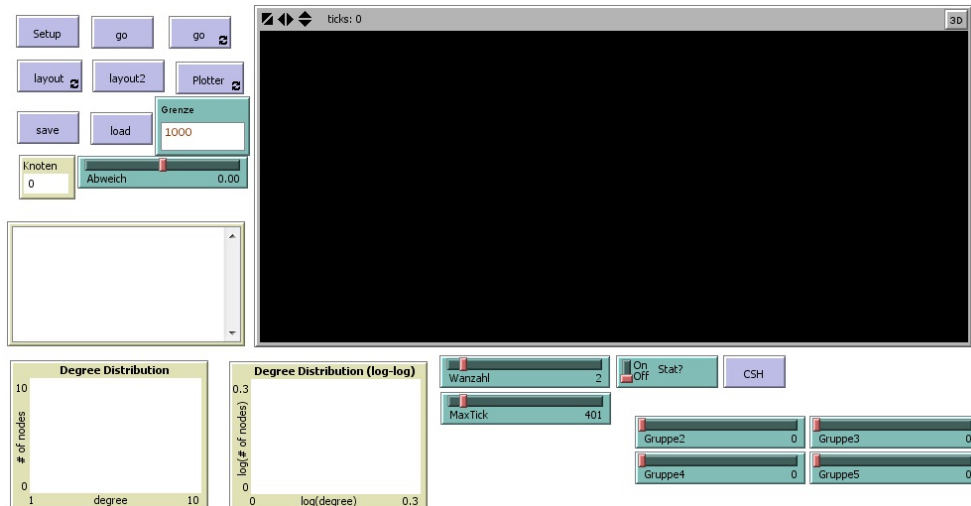


Abbildung 40: Das Tool zur Netzwerkerstellung

- 1) Als erstes muss mit dem Knopf "Setup" das Startdreieck erstellt werden
- 2) Nun können mit "Abweich" und "Grenze" die Randbedingungen für die Erstellung eingestellt werden.
- 3) Mit dem Knopf "Go" wird ein Durchgang der Erstellung durchgeführt. Mit dem Knopf "Go", der in der unteren rechten Ecke ein Wiederholungssymbol besitzt, wird die Erstellung solange durchgeführt bis der Knopf erneut betätigt wird.
- 4) Die "Layout" Knöpfe dienen der besseren Anordnung der Knoten im Darstellungsfenster.
- 5) Mit dem Knopf "Plotter" wird die lineare und die logarithmische Knotengradverteilung ausgegeben. Zusätzlich wird die prozentuale Anzahl der Knoten nach den in der Arbeit beschriebenen Gruppen und die absolute Anzahl der Knoten in den beiden höchsten Gruppen ausgegeben.
- 6) "Safe" und "Load" dienen dem Speichern und Laden des Netzwerkes. Dabei werden immer zuerst die Knoten und danach die Kanten verarbeitet. Zu beachten ist, dass das

Vorwort zu den Kurzanleitungen

- Fenster identisch ist, aber die Funktionalität unterschiedlich. (Beispielsweise warnt auch das Load Fenster vor dem überschreiben. Diese Warnung kann ignoriert werden.)
- 7) Die nach den Gruppen benannten Schieberegler und der Knopf "CSH" dienen der im Text vorgestellten Funktion "Create Social Hubs"
 - 8) Wird "Stat?" aktiviert, kann das Netzwerk so oft wie bei "Wanzahl" eingestellt und bis zu der Größe, die bei „Max-Tick“ eingestellt wird, erzeugt werden. Die durchschnittliche Knotengradverteilung in Gruppen wird anschließend ausgegeben. Dies dient der schnellen Überprüfung der Netzwerkergebnisse, wenn etwas an der Erstellung verändert wurde.

Kurzanleitung für das Tool zur Netzwerkbeschriftung

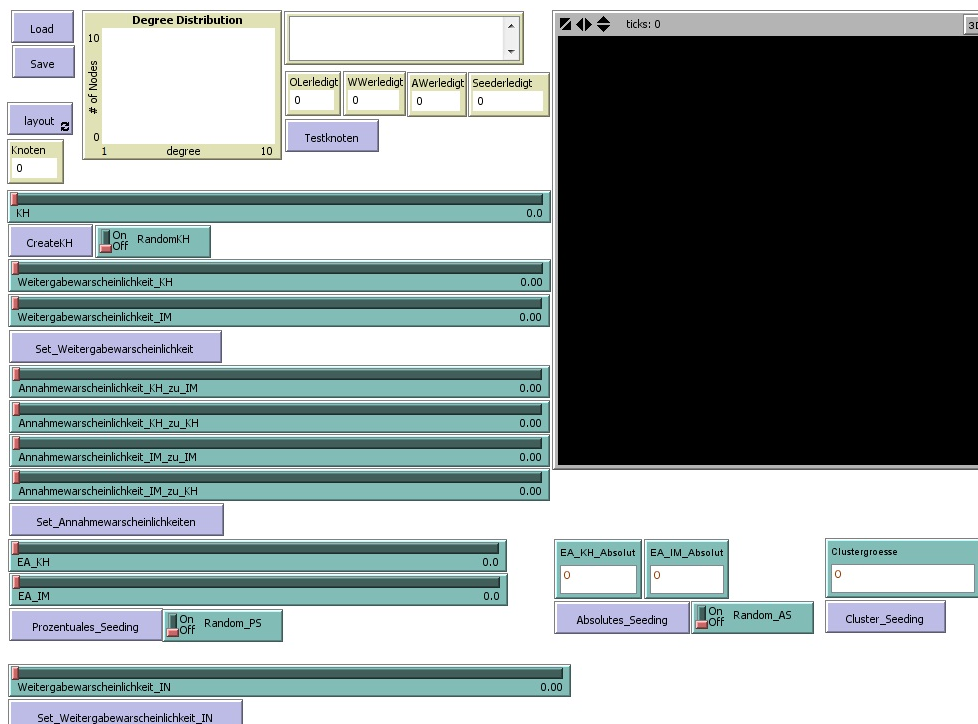


Abbildung 41: Das Tool zur Netzwerkbeschriftung

- 1) Nachdem das erstellte Netzwerk mit "Load" geladen wurde, kann mit der Beschriftung begonnen werden.
- 2) Mit "CreateKH" können, wie im Text beschrieben, die Knowledge Hubs im Netzwerk verteilt werden.
- 3) Danach können die mit den Schiebereglern eingestellten Wahrscheinlichkeiten mit "Set_Weitergabewahrscheinlichkeit" und "Set_Annahmewahrscheinlichkeit" in das Netzwerk übernommen werden.
- 4) Anschließend kann das Seeding mit den Knöpfen "Prozentuales_Seeding", "Absolutes_Seeding" und „Cluster_Seeding“ durchgeführt werden. Dazu wird die gewünschte Zahl an Innovators entweder prozentual mit den Schiebereglern oder in absoluten Werten in die Eingabefelder eingegeben.
- 5) Als letztes kann die Weitergabewahrscheinlichkeit der erstellten Innovators mit "Set_Weitergabewahrscheinlichkeit_IN" beeinflusst werden.
- 6) In den Fenstern "KHerledigt", "WWerledigt", "AWerledigt" und "Seederledigt" wird eine 1 angezeigt, sobald die Schritte zur Verteilung von Knowledge Hubs, Weitergabewahrscheinlichkeiten, Annahmewahrscheinlichkeiten und Innovators durchgeführt wurden.
- 7) Mit dem Knopf "Testknoten" kann ein zufällig ausgewählter Knoten und dessen Parameter angezeigt werden.
- 8) Abschließend muss die neue Knotendatei mit "Safe" für die Simulation gespeichert werden.

Kurzanleitung für das Tool zur Simulationsdurchführung

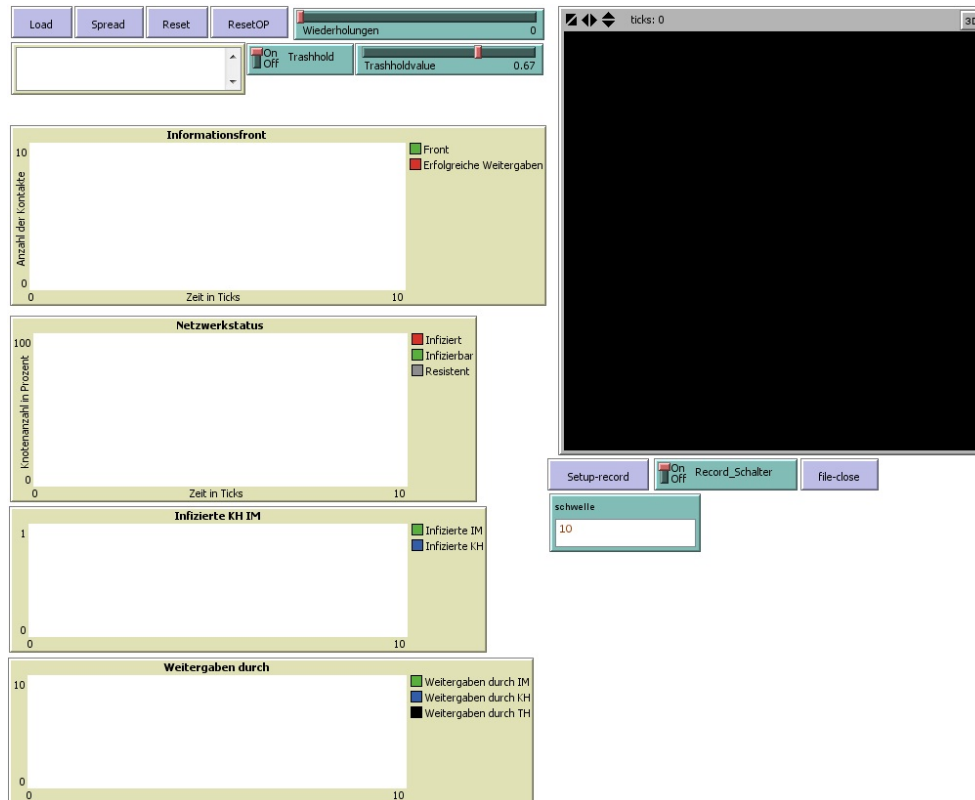


Abbildung 42: Das Tool zur Simulationsdurchführung

- 1) Nachdem die Knoten- und Kantendatei mit "Load" geladen wurde, kann mit "Spread" die Simulation so oft durchgeführt werden wie dies bei "Wiederholungen" eingestellt wurde.
- 2) Mit "Reset" kann die Simulation in den Ausgangszustand versetzt werden und mit "ResetOP" werden die Ausgabefenster gelöscht.
- 3) Mit "Threshold" kann der mit dem Schieberegler eingestellte Threshold für die Simulation aktiviert werden.
- 4) Mit "Schwelle" kann ein Schwellenwert der Ausbreitungsgröße in Prozent eingestellt werden ab dem zusätzliche Ausgaben erfolgen sollen. So wird die

Ausbreitungsanzahl (CC - Cascade Counter), die durchschnittliche Ausbreitungsgröße (CS - Cascade Size) und die durchschnittliche Ausbreitungsdauer (CD - Cascade Duration) ausgegeben.

- 5) Soll eine Auswertung mittels eines Tabellenkalkulationsprogrammes erfolgen, muss der "Record_Schalter" auf "on" gestellt werden und mit "Setup-record" eine Ausgabedatei zugewiesen werden. Nun werden alle Daten die auch in den Grafen dargestellt werden ebenfalls in die Datei geschrieben. Abschließend muss die Datei mit "file-close" geschlossen werden. Dabei werden folgende Werte ausgegeben: Tick (Zeitpunkt), Front (Anzahl der möglichen Weitergabepartner), W (Weitergaben), WIM (Weitergaben durch Imitators), WKH (Weitergaben durch Knowledge Hubs), WTH (Weitergaben durch Threshold), Ib (infizierbare Knoten), I (infizierte Knoten), R (resistente Knoten), IIM (infizierte Imitators) und IKH (infizierte Knowledge Hubs). Als Spaltentrenner wird das Semikolon benutzt.

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Alle Teile, die wörtlich oder sinngemäß einer Veröffentlichung entstammen, sind als solche kenntlich gemacht.

Die Arbeit wurde noch nicht veröffentlicht oder einer anderen Prüfungsbehörde vorgelegt.

Mittweida 27.8.2010

Steve Nebel, Unterschrift des Verfassers